

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 11
||:|||||

DB 200 LAHDDIPYEE 210

RESULT 8

P91.15 protein - Arabidopsis thaliana
B86286
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: B86286
R:Phelopsis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chou, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86286
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1451 <STO>
A:Cross-references: GB:AB005172; NID:95103820; PIDN:AAU39650.1; GSPDB:GN00141
A:Map position: 1

Query Match .7676%; Score 36; DB 2; Length 1451;
Best Local Similarity 54.5%; Pred. No. 41;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 11
||:|||||

DB 491 LSNEISVYDK 501

RESULT 9

W6WU31
E6 protein - human papillomavirus type 31
C:Species: human papillomavirus type 31
A:Note: host homo sapiens (man)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A32444
R:Goldborough, M.D.; Disilvestre, D.; Temple, G.F.; Lofincz, A.T.
Virology 171, 306-311, 1989
A>Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associated
A:Reference number: A94396; MUID:89299478; PMID:2545036
A:Accession: A32444
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-149 <GOLD>
A:Cross-references: GB:J04353; NID:q333048; PIDN:AAA6550.1; PID:9459916
C:Comment: This protein may be involved in the oncogenic potential of this virus.
C:Superfamily: papillomavirus E6 protein
C:Keywords: DNA binding; early protein; zinc finger
F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif

Query Match 74.5%; Score 35; DB 1; Length 149;
Best Local Similarity 63.6%; Pred. No. 5.1;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 11
||:|||||

DB 15 LSXELXPYEE 25

RESULT 10

JC7337
thiazole biosynthetic enzyme - Aspergillus oryzae
N:Alternate names: pyrithiamin resistance protein
C:Species: Aspergillus oryzae
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 03-Aug-2001
C:Accession: JC7337
R:Kubodera, T.; Yamashita, N.; Nishimura, A.
Biosci. Biotechnol. Biochem. 64, 1416-1421, 2000
A>Title: Pyrithiamin resistance gene (pirt) of Aspergillus oryzae: Cloning, character
A:Reference number: JC7337; MUID:20393935; PMID:10945258
A:Accession: JC7337
A:Molecule type: DNA
A:Residues: 1-327 <KUB>
A:Cross-references: GB:AF217503; PID:96694866; PIDN:AAF25444.1; PID:96694867
A:Experimental source: strain wild type
C:Genetics:

Query Match 74.5%; Score 35; DB 2; Length 327;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELXIPYEE 11
||:|||||

DB 138 ELXIPYEE 145

RESULT 11

C70133
conserved hypothetical protein BB0267 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: C70133
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Laibinra, R.; M
son, D.; Peterson, J.; Kellavag, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; W
Bowman, C.; Garland, S.; Fujii, C.; Colton, M.D.; Horst, K.; Roberts, K.; Hatch, B
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: C70133
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-634 <KIM>
A:Cross-references: GB:AF001137; GB:AF000783; NID:g2688160; PIDN:AA66681.1; PID:g26
A:Experimental source: strain B31

Query Match 74.5%; Score 35; DB 2; Length 634;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SXELXPYEE 11
||:|||||

DB 43 SLEISVYKE 52

RESULT 12

T30452
hypothetical protein ORF102 - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LDMNV
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30452
R:Kusio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; I
Virology 253, 17-34, 1999
A>Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantria
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30452
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-381 <KUZ>
A:Cross-references: EMBL:AF081810; NID:g3822234; PIDN:AACT02088.1; PID:g3822337

Query Match 72.3%; Score 34; DB 2; Length 381;
 Best Local Similarity 75.0%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 LSXELXPYEE 11
 ||:||||
 Db 41 ELDPYEE 48

RESULT 13

T27289
 hypothetical protein Y66A7A.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T27289
 R:Matthews, L.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20338
 A:Accession: T27289
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-738 <WIL>
 A:Cross-references: EMBL:AL032622; PIDN:CAA21499.1; GSPDB:GN00021; CESP:Y66A7A.5
 A:Experimental source: clone Y66A7A
 C:Genetics:
 A:Gene: CESP:Y66A7A.5
 A:Map position: 3
 A:introns: 201/3; 221/3; 287/1; 410/3; 551/1; 672/3; 723/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein Y66A7A.5

Query Match 72.3%; Score 34; DB 2; Length 738;
 Best Local Similarity 60.0%; Pred. No. 50;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 10
 ||:||||
 Db 673 LALEDPYEE 682

RESULT 14

MA0597
 DNA-directed DNA polymerase (EC 2.7.7.7) I - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 17-Mar-2000
 C:Accession: A40597; Y75363; S42112; S42235
 R:Gutman, P.D.; Fuchs, P.; Ouyang, L.; Minton, K.W.
 J. Bacteriol. 175, 3581-3590, 1993
 A:Title: Identification, sequencing, and targeted mutagenesis of a DNA polymerase gene I
 A:Reference number: A40597; MUID:932/3728; PMID:8501062
 A:Accession: A40597
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-921 <GUT>
 A:Cross-references: EMBL:L14581; NID:9290067; PIDN:AAC36974.1; PID:9290068
 A:Experimental source: strain R1
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Shen, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: E75363
 A:Molecule type: DNA
 A:Residues: 1-921 <GUT>
 A:Cross-references: EMBL:AF002012; GB:AE000513; NID:96459473; PIDN:AAF11264.1; PID:9645947
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRI707
 A:Map position: 1
 C:Superfamily: DNA-directed DNA polymerase I
 C:Keywords: nucleotidyltransferase

Query Match 72.3%; Score 34; DB 2; Length 921;
 Best Local Similarity 63.6%; Pred. No. 65;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 11
 ||:||||
 Db 768 LSNDLGPYEE 778

RESULT 15

AB1440
 hypothetical protein lln0057 [imported] - Listeria innocua (strain C1111262)
 C:Species: Listeria innocua
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AB1440
 R:Glaser, P.; Franke, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlen, K.D.; Fsiht,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kuntz, F.; Kurapkai, G.; Madueno, E.; Maltournam, A.;
 Ok, C.; Schlueter, T.; Stoces, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, R.; Wehla
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AB1440
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-233 <GUA>
 A:Cross-references: GB:AL592022; PIDN:CAC95290.1; PID:q16412478; GSPDB:CN00178
 A:Experimental source: strain C1111262
 C:Genetics:
 A:Gene: lln0057

Query Match 70.2%; Score 33; DB 2; Length 233;
 Best Local Similarity 63.6%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 11
 ||:||||
 Db 221 LFDLEDPYEE 231

Search completed: January 15, 2003, 12:37:13
 Job time: 19.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 Seconds
(without alignments)
72.584 Million cell updates/sec

Title: US-09-823-649A-2
Perfect score: 47
Sequence: 1 LSXELXIPYEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	93.6	831	1 DPOE_THERH	P30313 thermus the
2	44	93.6	832	1 DPOI_THERAQ	P19821 thermus agu
3	44	93.6	834	1 DPOI_THERCA	P01914 thermus cal
4	44	93.6	834	1 DPOI_THERTH	P52028 thermus the
5	38	80.9	229	1 CG18_YEAST	P18794 saccharomyc
6	35	74.5	149	1 VEB6_HPV31	P17386 human papil
7	35	74.5	327	1 TH14_ASPOR	P52027 deinoceccus
8	34	72.3	956	1 DPOI_DTRRA	P22318 saccharomyc
9	33	70.2	326	1 TH14_YEAST	P52318 saccharomyc
10	33	70.2	390	1 MOEA_STYV3	P53568 synchocyst
11	33	70.2	433	1 Y103_METUA	O57557 methanococ
12	33	70.2	499	1 C1W5_HUMAN	O55279 homo sapien
13	33	70.2	1136	1 POLG_OMV	P20234 o genome po
14	32	68.1	197	1 Y928_METUA	O58338 methanococ
15	32	68.1	227	1 IDI1_HUMAN	O13907 homo sapien
16	32	68.1	227	1 IDI1_MESAU	O35586 mesocricetu
17	32	68.1	227	1 IDI1_MOUSE	P58044 mus musculu
18	32	68.1	227	1 IDI1_RAT	O35760 rattus norv
19	32	68.1	753	1 NECL_MOUSE	P21662 mus musculu
20	32	68.1	2517	1 NCR2_HUMAN	O97618 h nuclear r
21	31	66.0	144	1 YG66_BPA18	O9154 bacterioph
22	31	66.0	162	1 PUR6_BACSU	P12044 bacillus su
23	31	66.0	208	1 Y103_ECOLI	P75805 escherichia
24	31	66.0	238	1 RM08_YEAST	P22353 saccharomyc
25	31	66.0	403	1 ACKA_HELPJ	O92405 heliocobacte
26	31	66.0	752	1 NECL_RAT	P28840 rattus norv
27	31	66.0	753	1 NECL_HUMAN	P29120 homo sapien
28	31	66.0	833	1 DPOI_THERFI	O52225 thermus fil
29	31	66.0	1167	1 C1AA_BACTO	P56956 bacillus th
30	30	63.8	266	1 AKOE_HELPJ	O92438 heliocobacte
31	30	63.8	274	1 TH14_AERPE	O97920 aeropyrum p
32	30	63.8	285	1 WTAF_CABEL	O09438 caenorhabdi
33	30	63.8	326	1 TMOF_PSEME	O03304 pseudomonas

ALIGNMENTS

RESULT 1	ID	DPOF_THERH	STANDARD:	PRT:	831 AA.
AC	P30313:				
DI	01-APR-1993 (rel. 25, Created)				
DI	01-APR-1993 (rel. 25, Last sequence update)				
DI	15-JUN-2002 (rel. 41, Last annotation update)				
DE	DNA polymerase 1, thermostable (EC 2.7.7.7) (Tf1 polymerase 1).				
GN	POLA OR POL.				
OS	Thermus thermophilus.				
OC	Bacteriia; Thermus/Deinococcus group; Deinococci; Thermales;				
OC	Thermaceae; Thermus.				
OX	NCBI_TaxID=274;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ACM B-1257;				
RX	MEDLINE=93087201; PubMed=1454544;				
RA	AKmetzjanov A.A., Vakhitov V.A.;				
RT	"Molecular cloning and nucleotide sequence of the DNA polymerase gene				
RT	from Thermus flavus.";				
RL	Nucleic Acids Res. 20:5839-5839(1992).				
CC	-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + [DNA](N).				
CC	-1- BIOTECNOLOGY: Used in the PCR method because of its high				
CC	thermostability.				
CC	-1- MISCELLANEOUS: Has an optimal temperature above 70 degrees				
CC	Celsius; can be used at temperatures up to 95 degree Celsius.				
CC	-1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: X66105; CAA6900.1; -				
DR	PIR: S24929; S24929.				
DR	PIR: S26675; S26675.				
DR	HSSP: P19821; ITAQ.				
DR	InterPro: IPR002421; 5_3_exonuclease.				
DR	InterPro: IPR001098; DNA pol.				
DR	InterPro: IPR002298; DNA pol.				
DR	InterPro: IPR000513; EXO_N.I.				
DR	InterPro: IPR003583; HNH_1.				
DR	InterPro: IPR003584; HNH_2.				
DR	InterPro: IPR001532; XPC_H_Rad.				
DR	Pfam: PF00476; DNA_pol_A; 1.				
DR	Pfam: PF021367; 5_3_exonuclease; 1.				
DR	Pfam: PF02739; 5_3_exonuclease; 1.				
DR	PRINTS: PRO0866; DNAPO1.				
DR	SMART: SM00475; 53EXOC; 1.				
DR	SMART: SM00278; HNH1; 1.				
DR	SMART: SM00279; HNH2; 1.				
DR	SMART: SM00482; POLAC; 1.				

DR SMART; SM00485; XPGN; 1.
 DR TIGR00593; polA; 1.
 DR PROSITE; PS00447; DNA.POLYMERASE.A; 1.
 KM Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
 KM DNA-binding; 409
 SQ DOMAIN 831 POLYMERASE (BY SIMILARITY).
 SQ SEQUENCE 831 AA; 93783 MW; 96FP93CEFA3CA536D CRC64;
 Query Match 93.6%; Score 44; DB 1; Length 831;
 Best Local Similarity 81.8%; Pred. No. 0.19;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSXELX1PYEE 11
 DB 677 LSSELISPYEE 687
 RESULT 2
 DPOL_THEADQ STANDARD; PRT; 832 AA.
 ID DPOL_THEADQ
 AC P19821;
 DT 01-FEB-1991 (rel. 17, Created)
 DT 01-FEB-1991 (rel. 17, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE DNA polymerase I, thermostable (EC 2.7.7.7) (Tag polymerase 1).
 GN POLA OR POLI.
 OS Thermus aquaticus.
 OC Bacteria; Thermus/Delinoococcus group; Delinococci; Thermales;
 OC Thermaceae; Thermus.
 CX NCBI_TaxID=271;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89197950; PubMed-2649500;
 RA Lawyer F.C., Stoffel S., Saiki R.K., Myambo K., Drummond R.,
 RA Gelfand D.H.;
 RT "Isolation, characterization, and expression in Escherichia coli of
 RT the DNA polymerase gene from Thermus aquaticus.";
 RL J. Biol. Chem. 264:6427-6437(1989).
 RN 12
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-YT1;
 RX MEDLINE-95204371; PubMed-7896728;
 RA Ishino Y., Ueno T., Miyagi M., Demori T., Imanura M., Tsunasawa S.,
 RA Kato I.;
 RT "Overproduction of Thermus aquaticus DNA polymerase and its
 RT structural analysis by ion-spray mass spectrometry.";
 RL J. Biochem. 116:1019-1024(1994).
 RN 13
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE-95364959; PubMed-7637814;
 RA Kim Y., Eom S.H., Wang J., Lee D.-S., Suh S.W., Stellz T.A.;
 RT "Crystal structure of Thermus aquaticus DNA polymerase.";
 RL Nature 376:612-616(1995).
 RN 14
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 290-832.
 RX MEDLINE-96016150; PubMed-7568114;
 RA Korolev S., Nayal M., Barnes W.M., di Cera E., Waksman G.;
 RT "Crystal structure of the large fragment of Thermus aquaticus DNA
 RT polymerase I at 2.5-A resolution: structural basis for
 RT thermostability.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9264-9268(1995).
 RN 15
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE-96533982; PubMed-8717047;
 RA Eom S.H., Wang J., Stellz T.A.;
 RT "Structure of Tag polymerase with DNA at the polymerase active site.";
 RL Nature 382:278-281(1996).
 RN 16
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 295-832.
 RX MEDLINE-9907817; PubMed-9857206;
 RA Li Y., Korolev S., Waksman G.;
 RT "Crystal structures of open and closed forms of binary and ternary
 RT complexes of the large fragment of Thermus aquaticus DNA polymerase

RT 1: structural basis for nucleotide incorporation.";
 RL EMBO J. 17:7514-7525(1998).
 RN 17
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 290-832.
 RX MEDLINE-98266352; PubMed-9605316;
 RA Li Y., Kong Y., Korolev S., Waksman G.;
 RT "Crystal structures of the Klenow fragment of Thermus aquaticus DNA
 RT polymerase I complexed with deoxyribonucleoside triphosphates.";
 RL Protein Sci. 7:1116-1123(1998).
 RN 18
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE-98445410; PubMed-9770525;
 RA Morali R., Sharkey D.J., Dais J.L., Murthy H.M.K.;
 RT "Crystal structure of Tag DNA polymerase in complex with an inhibitory
 RT pad: the pad is directed against an intermediate in the helix-coil
 RT dynamics of the enzyme.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12562-12567(1998).
 RN 19
 RP X-RAY CRYSTALLOGRAPHY (2.30 ANGSTROMS) OF 293-831.
 RX MEDLINE-99380545; PubMed-10449720;
 RA Li Y., Milakov V., Waksman G.;
 RT "Structure-based design of Tag DNA polymerases with improved
 RT properties of dideoxynucleotide incorporation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9491-9496(1999).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA)(N).
 CC -1- BIOCHEMISTRY: Used in the PCR method because of its high
 CC thermostability. Has a relatively high error rate because it lacks
 CC exonuclease proofreading functionality.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J04639; AAA27507.1;
 DR EMBL: D32013; BAA06775.1;
 DR PIR: A33530; A33530.
 DR PDB: 1TQA; 07-DEC-96.
 DR PDB: 1PAU; 11-APR-97.
 DR PDB: 1KTO; 08-NOV-96.
 DR PDB: 2KTO; 22-DEC-99.
 DR PDB: 4KTO; 22-DEC-99.
 DR PDB: 4KTO; 22-DEC-99.
 DR PDB: 5KTO; 30-SEP-98.
 DR PDB: 1HCX; 14-OCT-98.
 DR PDB: 1UXE; 12-MAR-97.
 DR PDB: 1OSS; 16-AUG-99.
 DR PDB: 1OSY; 16-AUG-99.
 DR PDB: 1OTM; 16-AUG-99.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR000513; Exo_N1.
 DR InterPro: IPR003583; HHH_1.
 DR InterPro: IPR003584; HHH_2.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuc; 1.
 DR PRINTS: PR00868; DNAPOL1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00278; Hhh1; 1.
 DR SMART: SM00279; Hhh2; 1.
 DR SMART: SM00482; PolA; 1.
 DR TIGR00593; polA; 1.
 DR TRANSFER: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA.POLYMERASE.A; 1.
 KM Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
 KM DNA-binding; 3D-structure.
 FM DOMAIN 410 832 POLYMERASE (BY SIMILARITY).

FT CONFLICT 155 155 V -> A (IN REF. 2).
SQ SEQUENCE 832 AA; 93909 MW; F1731055B5246F03 CRC64;
Query Match 93.6%; Score 44; DB 1; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 LSXELXIPYEE 11
DB 678 LSQELAIYEE 688

RESULT 3
DPOL_THRCA STANDARD; PRT; 834 AA.
AC P80194;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (TAC polymerase 1).
GN POLA.
OS Thermus caldophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GK24;
RA Kwon S.-T., Kim J.S., Park J.H., Kim H., Lee D.-S.;
RN Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
RT 121
RP SEQUENCE OF 1-21.
RC STRAIN-GK24;
RA MEDLINE-93285135; PubMed=8508785;
RT Purification and characterization of Thermus caldophilus GK24 DNA polymerase.
CC Eur. J. Biochem. 214:135-140(1993).
CC -1- FUNCTION: THERMOSTABLE ENZYME THAT HAS 5'-TO-3' EXONUCLEASE
CC ACTIVITY AND NO 3'-TO-5' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D62584; AAB81398.1; -
CC PIR: S33287; S33287.
DR HSSP: P19821; ITNO.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR00513; Exo_N.I.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF00476; DNA_pol_A; 1.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuclease; 1.
DR PRINTS: PR00868; DNAPOLI.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; Hhh1; 1.
DR SMART: SM00279; Hhh2; 1.
DR SMART: SM00482; POLAC; 1.
DR TIGRPFAM: TIGR00593; pola; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.

FT DOMAIN 412 834 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 834 AA; 93798 MW; A851F93C3076346E CRC64;
Query Match 93.6%; Score 44; DB 1; Length 834;
Best Local Similarity 81.8%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 LSXELXIPYEE 11
DB 680 LSQELAIYEE 690

RESULT 4
DPOL_THRTH STANDARD; PRT; 834 AA.
AC P52028;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (Tth polymerase 1).
GN POLA.
OS Thermus thermophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HB8 / ATCC 27634;
RA Asakura K., Komatsubara H., Soga S., Yomo T., Oka M., Eml S.,
RA Urabe I.;
RT Cloning, nucleotide sequence, and expression in Escherichia coli of
RT DNA polymerase gene (pola) from Thermus thermophilus HB8.
RT J. Ferment. Bioeng. 76:265-269(1993).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D28878; BAA0603.1; -
CC HSSP: P19821; ITNO.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR00513; Exo_N.I.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF00476; DNA_pol_A; 1.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuclease; 1.
DR PRINTS: PR00868; DNAPOLI.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; Hhh1; 1.
DR SMART: SM00279; Hhh2; 1.
DR SMART: SM00482; POLAC; 1.
DR TIGRPFAM: TIGR00593; pola; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.

```

CC NCBI_TaxID=10585;
CC [1]
CC SEQUENCE FROM N.A.
CC Middle=8929478; PubMed=2545036;
CC Goldborough M.D., Distlvestre D., "Temple G.F., Lorincz A.".;
CC "Nucleotide sequence of human papillomavirus type 31: a cervical
CC RT virology 171:306-311(1989)."
CC FL
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL
CC OF THIS VIRUS (CERVICAL NEOPLASIA-ASSOCIATED VIRUS) .
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO) .
CC -1- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J04353; AAA46950.1; -.
CC DR PIR: A32444; W6WLJ1.
CC DR InterPro: IPR001334; E6.
CC DR Pfam: PF00518; E6; 1.
CC KM Early protein; DNA-binding; Nuclear protein; zinc-finger;
CC KW Oncogene.
CC FT ZN_FING 30 66 POTENTIAL.
CC FT FT 103 139 POTENTIAL.
CC SQ SEQUENCE 149 AA; 17713 MW; 61D2A86C362767D CRC64;
CC -----
CC Query Match 74.5%; Score 35; DR 1; Length 149;
CC Best Local Similarity 63.6%; Pred. No. 2.4;
CC Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC QY 1 LSXELXIPYEE 11 -
CC Db 15 LSSALEIPYDE 25
CC -----
CC RESULT 7
CC TH14_ASPOR
CC ID TH14_ASPOR STANDARD; PRT; 327 AA.
CC AC Q9U0Z9;
CC DT 16-OCT-2001 (Ref. 40, Created)
CC DT 16-OCT-2001 (Ref. 40, Last sequence update)
CC DT 16-OCT-2001 (Ref. 40, Last annotation update)
CC DE Thiazole biosynthetic enzyme, mitochondrial precursor.
CC GN TH14 OR PTFA.
CC OS Aspergillus oryzae.
CC OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
CC OC Eucristales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
CC XX NCBI_TaxID=5062;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=HL1134;
CC MEDLINE=20399355; PubMed=10945258;
CC Kubodera T., Yamashita N., Nishimura A.;
CC "Pythianine resistance gene (prrA) of Aspergillus oryzae: cloning,
CC transformation and application as a dominant selectable marker for
CC transformation.";
CC RL Biosci. Biotechnol. Biochem. 64:1416-1421(2000).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF THE THIAMINE PRECURSOR
CC THIAZOLE (BY SIMILARITY) .
CC -1- SUBCELLULAR LOCATION: Mitochondrial (Potential).
CC -1- SIMILARITY: BELONGS TO THE TH14 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF217503; AAF25444.1;
DR InterPro: IPR002922; Thi4.
DR Pfam: PF01946; Thi4; 1.
DR TIGRPFAMs: TIGR00292; Thi4; 1.
KW Thiamine biosynthesis; Mitochondrion; Transl. peptid; FAD; NAD.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 327 THIAZOLE BIOSYNTHETIC ENZYME.
FT NP BIND 77 107 FAD OR NAD (POTENTIAL).
SQ SEQUENCE 327 AA; 35099 MW; 7C561EB06742B2AE CRC64;

Query Match
Best Local Similarity 74.5%; Score 35; DB 1; Length 327;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELIXPER 11
DB 138 ELGVPYEE 145

RESULT 8
DPOL_DEIRA STANDARD; PRT; 956 AA.
AC P52027;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL I).
GN POLA OR DR1707.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE OF 36-956 FROM N.A.
RC STRAIN=RI;
RA MEDLINE=93273728; PubMed=8501062;
RA Gutman P.D., Fuchs P., Ouyang L., Minton K.W.;
RT Identification, sequencing, and targeted mutagenesis of a DNA
RT polymerase gene required for the extreme radioresistance of
RT Deinococcus radiodurans";
RL J. Bacteriol. 175:3581-3590(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RA MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.D., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.B.,
RA Moffat K.S., Qin B., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarewicz K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans RI.";
RL Science 286:1571-1577(1999).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA](N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: L14581; AAC36974.1;
DR EMBL: AE002012; AAF11264.1;
DR HSSP: P19821; 2KTQ.
DR TIGR: DR1707;
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR003583; HNH_1.
DR InterPro: IPR003584; HNH_2.
DR Pfam: PF00476; DNA_Pol_A_1.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuc_N; 1.
DR PRINTS: PR00688; DNAPOL1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; HNH1; 1.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00482; POLA; 1.
DR TIGRPFAMs: TIGR00593; polA; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
FT CONFLICT 48 46 T -> A (IN REF. 1).
FT CONFLICT 77 77 H -> D (IN REF. 1).
FT CONFLICT 108 108 P -> L (IN REF. 1).
FT CONFLICT 128 128 P -> R (IN REF. 1).
FT CONFLICT 406 406 P -> R (IN REF. 1).
FT CONFLICT 540 540 L -> V (IN REF. 1).
SQ SEQUENCE 956 AA; 105659 MW; 6ABHP117D75AB84A CRC64;

Query Match
Best Local Similarity 72.3%; Score 34; DB 1; Length 956;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ISXELXPYEE 11
DB 803 LSNDIGIPYAE 813

RESULT 9
TH14_YEAST
ID TH14_YEAST STANDARD; PRT; 326 AA.
AC P32318;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thiazole biosynthetic enzyme, mitochondrial precursor.
GN TH14 OR MO11 OR WS935 OR YNR144W OR G6620.
OS Saccharomyces cerevisiae (Haker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93070608; PubMed=1441749;
RA Prækel U.M., Meacock P.A.;
RT "MO11, a Saccharomyces cerevisiae gene that is highly expressed in
RT early stationary phase during growth on molasses.";
RL Yeast 8:699-710(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=96158062; PubMed=8585325;
RA Skala J., Nawrocki A., Goffeau A.;
RT "The sequence of a 27 kb segment on the right arm of chromosome VII
RT from Saccharomyces cerevisiae reveals MO11, NAT2, RPL30B, NSR1, CYS4,
RT PEM1/CMO2, NSR1 genes and ten new open reading frames.";

```

```

RL Yeast 11:1421-1427(1995).
RN [4]
RP THIAMINE REGULATION.
RX MEDLINE=95028146; PubMed=7941734;
RA Praekelt U.M., Byrne K.L., Meacock P.A.;
RT Regulation of Th14 (MOL1), a thiamine-biosynthetic gene of
RL Saccharomyces cerevisiae.
RN Yeast 10:481-490(1994).
RN [5]
RP FUNCTION IN DNA DAMAGE TOLERANCE.
RX MEDLINE=96035046; PubMed=9367751;
RA Machado C.R., Praekelt U.M., de Oliveira R.C., Barbosa A.C.,
RA Byrne K.L., Meacock P.A., Menck C.F.;
RT Dual role for the yeast Th14 gene in thiamine biosynthesis and DNA
RT damage tolerance.
RL J. Mol. Biol. 273:114-121(1997).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF THE THIAMINE PRECURSOR
CC THIAZOLE. ALSO SEEMS TO HAVE A ROLE IN MITOCHONDRIAL DNA DAMAGE
CC TOLERANCE.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (Potential).
CC -1- INDUCTION: RERESSED BY THIAMINE.
CC -1- MISCELLANEOUS: EXPRESSED AT HIGH LEVELS IN THE EARLY STATIONARY
CC PHASE OF BATCH CULTURES GROWING ON MOLASSES, AN INDUSTRIAL MEDIUM.
CC -1- SIMILARITY: BELONGS TO THE TH14 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X61669; CAA3843.1; -
DR EMBL: X85607; CA59802.1; -
DR EMBL: Z72829; CA97157.1; -
DR EMBL: Z72930; CA97159.1; -
DR PIR: S25321; S25321.
DR SCD: S0003376; TH14.
DR InterPro: IPR002922; Th14.
DR Pfam: PF01946; Th14; 1.
DR TIGRFS: TIGR00292; Th14; 1.
KW Thiamine biosynthesis; Mitochondrion; Transit peptide; FAD: NMD.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 326 THIAZOLE BIOSYNTHETIC ENZYME.
FT NP BIND 67 97 FAD OR NAD (POTENTIAL).
SQ SEQUENCE 326 AA; 34991 MW; 8437902CE0DF02 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 326;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELIXIRYE 11
DB 128 ELIXIRYE 135

RESULT 10
MOEA_SYNY3
AC Q55368; STANDARD; PRT; 390 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Molybdopterin biosynthesis protein moea.
GN MOEA OR STR0900.
OS Synchocystis sp. (Strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;

```

```

RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. 1. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome."
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF A DEMOLYBDO-COPACTOR
CC (MOLYBDOPTERIN). NECESSARY FOR MOLYBDO-ENZYMES (BY SIMILARITY).
CC -1- PATHWAY: MOLYBDENUM COPACTOR BIOSYNTHESIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE C-TERMINAL OR CINNAMON/GERANYL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64003; BAA10450.1; -
DR InterPro: IPR001453; MOCF_biosynth.
DR InterPro: IPR005111; Moea_C.
DR InterPro: IPR005110; Moea_N.
DR Pfam: PF00994; MOCF_biosynth; 1.
DR Pfam: PF03453; Moea_N; 1.
DR Pfam: PF03454; Moea_C; 1.
DR Pfam: PD002460; MOCF_biosynth; 1.
DR Probab: PD002460; MOCF_biosynth; 1.
DR TIGRFS: TIGR00177; molyb_syn; 1.
DR PROSITE: PS01079; MOCF_BIOSYNTHESIS_2; 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 390 AA; 43394 MW; 848183B610CA14 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 390;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELIXIRYE 11
DB 103 DUXIRYE 110

RESULT 11
Y103_METUA
ID Y103_METUA STANDARD; PRT; 433 AA.
AC Q57567;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M10103.
GN M10103.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake O., Adams J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.P., Fuhmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna K.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: BELONGS TO THE UPF0063 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U67467; AAB98083.1; -
 DR TIGR; M20103; -
 DR InterPro; IPR000385; MoaA_NiFB_PpGE.
 DR Pfam; PF01444; MoaA_NiFB_PpGE; 1.
 KM Hypothetical protein: Complete proteome.
 SQ SEQUENCE 433 AA; 49866 MW; F82576531D12142 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 433;
 Best Local Similarity 75.0%; Pred. No. 19;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELXIPYEE 11
 1:|||||
 DB 33 EVELPYEE 40

RESULT 12
 C1M5_HUMAN STANDARD; PRT; 499 AA.
 ID C1M5_HUMAN
 AC 095279;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Potassium channel subfamily K member 5 (Acid-sensitive potassium
 DE channel protein TASK-2) (TWIK-related acid-sensitive K⁺ channel 2).
 GN KCNK5 OR TASK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=99030343; Pubmed=9812978;
 RA Reyes R., Duprat F., Lesage F., Fink M., Salinas M., Farman N.,
 RA Lazdunski M.;
 RT Cloning and expression of a novel pH-sensitive two pore domain K⁺
 RT channel from human kidney.
 RL J. Biol. Chem. 273:30863-30869(1998).
 CC -1- FUNCTION: pH DEPENDENT, VOLTAGE INSENSITIVE. OUTWARDLY RECTIFYING
 CC POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS LOST AT HIGH EXTERNAL
 CC K⁺ CONCENTRATIONS.
 CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: ABUNDANT EXPRESSION IN KIDNEY, ALSO DETECTED
 CC IN LIVER, PLACENTA AND SMALL INTESTINE. IN THE KIDNEY, EXPRESSION
 CC IS RESTRICTED TO THE DISTAL TUBULES AND COLLECTING DUCTS. NOT
 CC EXPRESSED IN PROXIMAL TUBULES OR GLOMERULI.
 CC -1- MISCELLANEOUS: INHIBITED BY QUININE, QUINIDINE AND EXTERNAL
 CC ACIDIFICATION.
 CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
 CC CHANNELS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF084830; AAC79458.1; -
 DR Genbank; HGNC:6280; KCNK5.
 DR KIM; 603493; -
 DR InterPro; IPR003280; K+channel_2pore.
 DR InterPro; IPR001622; K+channel_pore.
 DR InterPro; IPR000636; M+channel_nig.

DR Pfam; PF00520; Ion_trans; 1.
 DR PRINTS; PF01333; 2PORKCHANNEL.
 KM Ionic channel; Transmembrane; Ion transport; Potassium transport;
 KW Glycoprotein.

FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 8 26 POTENTIAL.
 FT DOMAIN 85 112 PORE-FORMING 1 (POTENTIAL).
 FT TRANSMEM 113 133 POTENTIAL.
 FT DOMAIN 134 157 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 158 180 POTENTIAL.
 FT DOMAIN 190 215 PORE-FORMING 2 (POTENTIAL).
 FT TRANSMEM 230 250 POTENTIAL.
 FT DOMAIN 251 325 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 499 AA; 55130 MW; E871A7A4823DDA00 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 499;
 Best Local Similarity 62.5%; Pred. No. 22;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELXIPYEE 11
 1:|||||
 DB 478 ELSTPYEE 485

RESULT 13
 POLG_OMV STANDARD; PRT; 1136 AA.
 ID POLG_OMV
 AC P20234;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide [contains: Nuclear inclusion protein A (NI-A) (NIA)
 DE (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear inclusion
 DE protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC 2.7.7.48);
 DE Coat protein (CP) (Fragment).
 OS Ornithogalum moscatis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
 OC Polyvirus.
 ON NCBI_Taxid=12204;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Burger J.T., Brand R.J., Rybicki E.P.;
 RL Submitted (NOV-1990) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: NI-NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
 CC further restricted by preferences for the amino acids in p6 - p1'
 CC that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
 CC Glu (Ser or Gly) for the enzyme from tobacco etch virus. The
 CC natural substrate is the viral polyprotein, but other proteins and
 CC oligopeptides containing the appropriate consensus sequence are
 CC also cleaved.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC {RNA}(N).
 CC -1- PFM: THE VIRAL RNA OF POLYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POLYVIRUSES POLYPEPTIDE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D00615; BAA00490.1; -
 DR PIR; J00494; J00494.
 DR MEROPS; C04.005; -
 DR InterPro; IPR001730; peptidase_C4.

DR InterPro: IPR001592; Pcty_coat.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR Pfam: PF00767; Pcty_coat_1.
 DR Pfam: PF00863; Peptidase_C4_1.
 KM Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
 KM Coat protein; Polyprotein.
 FT NON_TER 1
 FT CHAIN <1 365 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 366 883 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 884 1136 COAT PROTEIN.
 SQ SEQUENCE 1136 AA; 128914 MW; 0A7AE0527743FD61 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 1136;
 Best Local Similarity 75.0%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 ELXLYPEE 11
 : | | | |
 Db 698 KLSIPEE 705

RESULT 14
 ID Y928 METUA STANDARD; PRT; 197 AA.
 AC 058338:
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein Mj0928 (M.JhaHemK).
 GN Mj0928.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=9637999; PubMed=8689087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb L.-F., Adams M., Reich C.L.,
 RA Overbeek R., Kirkness E.F., Weislock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fultman J.L., Nguyen D.,
 RA Ultebert T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Kleck H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: BELONGS TO THE HEMK FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U67536; AAB98930.1; -.
 CC DR HSSP: Q58292; IDUS.
 CC DR TIGR: Mj0928; -.
 CC DR REBASE: 4480; M.JhaHemK.
 DR InterPro: IPR004557; Hemk_rel_arch.
 DR InterPro: IPR000051; SAM_bind
 DR InterPro: IPR003402; Unk_Met10.
 DR Pfam: PF02475; Met_10; 1.
 DR TIGRFAMS: TIGR00537; hemk_rel_arch; 1.
 DR PROSITE: PS00092; N6_Mtase; UNKNOW1.
 KM Hypothetical protein; Transferase; Methyltransferase;
 KW Complete proteome.

SQ SEQUENCE 197 AA; 22057 MW; 2AE2377EE8EB211 CRC64;
 Query Match 68.1%; Score 32; DB 1; Length 197;
 Best Local Similarity 45.5%; Pred. No. 13;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 LSXELIPEE 11
 : | | | |
 Db 178 ISAKLYPEE 188

RESULT 15
 ID ID11 HUMAN STANDARD; PRT; 227 AA.
 AC Q13907:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Isopentenyl-diphosphate delta-isomerase 1 (PC 5.3.3.2) (IPP isomerase
 DE 1) (Isopentenyl pyrophosphate isomerase 1) (IPP11).
 GN ID11.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94292171; PubMed=8020941;
 RA Xuan J.W., Kowalski J., Chambers A.F., Dehardt D.T.;
 RT "A human promyelocyte mRNA transiently induced by TPA is homologous
 RT to yeast IPP isomerase.";
 RL Genomics 20:129-131(1994).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=96400329; PubMed=8806705;
 RA Hahn P.M., Xuan J.W., Chambers A.F., Poulter C.D.;
 RT "Human isopentenyl diphosphate: dimethylallyl diphosphate isomerase:
 RT overproduction, purification, and characterization.";
 RL Arch. Biochem. Biophys. 332:30-34(1996).
 CC -1- FUNCTION: CATALYZES THE 1,3-ALLYLIC REARRANGEMENT OF THE
 CC HOMOLALYLIC SUBSTRATE ISOPENTENYL (IPP) TO ITS HIGHLY
 CC ELECTROPHILIC ALLYLIC ISOMER, DIMETHYLLALLYL DIPHOSPHATE (DMAPP).
 CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate = dimethylallyl
 CC diphosphate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
 CC INCLUDE: ISOPRENOL BIOSYNTHETIC PATHWAY WHOSE END PRODUCTS
 CC CAROTENOIDS, RILE ACIDS AND CHOLESTEROL.
 CC -1- SUBCELLULAR LOCATION: Peroxisome.
 CC -1- SIMILARITY: BELONGS TO THE IPP ISOMERASE TYPE 1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X17025; CAA34890.1; ALU_INIT.
 CC DR Genew: HGNC:5387; ID11.
 CC DR MIM: 604055; -.
 DR InterPro: IPR002667; IPP_isomerase.
 DR InterPro: IPR000086; NUDIX_hydrolase.
 DR Pfam: PF00293; NUDIX_1
 DR Prodom: PD004109; IPP_isomerase; 1.
 KM Isomerase; Isoprene biosynthesis; Cholesterol biosynthesis;
 KM Sterol biosynthesis; Peroxisome; Magnesium.
 FT ACT_SITE 86 BY SIMILARITY.
 FT ACT_SITE 148 BY SIMILARITY.
 FT SITE 225 MICROBODY TARGETING SIGNAL.
 SQ SEQUENCE 227 AA; 26319 MW; 1255AC2C4D1B8D1 CRC64;

Query Match 68.1%; Score 32; DB 1; Length 227;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11

DB 112 LKAEIGIPLEE 122

Search completed: January 15, 2003, 11:21:19
Job time : 11.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 10:29:39 : Search time 25.5714 Seconds

(without alignments)
88.635 Million cell updates/sec

Title: US-09-823-649A-2

Perfect score: 7

Sequence: 1 LEXELXIPYEE 11

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	80.9	850	16	066691 aquifex aeo
2	37	78.7	474	16	080621 agrobacteri
3	36	76.6	1451	10	09X148 arbidopsis
4	35	74.5	634	16	044755 borrelia bu
5	35	74.5	1028	4	090052
6	35	74.5	1028	11	09JMB8 mus musculu
7	35	74.5	1028	11	P97528 ratius norv
8	34	72.3	274	5	094849 drosophila
9	34	72.3	291	16	08R889
10	34	72.3	381	12	09YMM5
11	34	72.3	738	5	09XWV5
12	33	70.2	233	16	092EP6
13	33	70.2	233	16	08VAQ3
14	33	70.2	461	16	0989F3
15	33	70.2	465	12	09Q4W5
16	33	70.2	502	11	09JKE2 mus musculu

17	33	70.2	597	3	059717	059717 schizosacch
18	33	70.2	692	17	097YM6	097YM6 sulfolobus
19	33	70.2	697	17	08RJ04	08RJ04 mehanosarc
20	33	70.2	1056	16	08RPF7	08RPF7 fusobacteri
21	33	70.2	4717	3	094248	094248 schizosacch
22	32	68.1	132	10	0920G2	0920G2 arbidopsis
23	32	68.1	153	16	0926Q7	0926Q7 listeria in
24	32	68.1	153	16	08Y314	08Y314 listeria mo
25	32	68.1	186	16	08U8F0	08U8F0 agrobacteri
26	32	68.1	199	4	09NMW6	09NMW6 homo sapien
27	32	68.1	228	4	0961Z4	0961Z4 homo sapien
28	32	68.1	228	4	08WUX8	08WUX8 homo sapien
29	32	68.1	228	4	09B074	09B074 homo sapien
30	32	68.1	346	2	093N02	093N02 escherichia
31	32	68.1	359	4	015268	015268 homo sapien
32	32	68.1	402	17	09YK31	09YK31 aetopyrum p
33	32	68.1	412	17	08Y7Z4	08Y7Z4 pyrococcus
34	32	68.1	417	17	09VIT1	09VIT1 pyrococcus
35	32	68.1	469	17	030094	030094 archaeoglob
36	32	68.1	627	5	08SVF8	08SVF8 encephalito
37	32	68.1	634	10	08S9X6	08S9X6 oryza sativ
38	32	68.1	1085	5	024363	024363 drosophila
39	32	68.1	1234	5	024690	024690 drosophila
40	32	68.1	1493	16	067512	067512 aquifex aeo
41	32	68.1	1890	10	049528	049528 arbidopsis
42	31	66.0	144	16	092FK3	092FK3 listeria in
43	31	66.0	144	16	08Y4X6	08Y4X6 listeria mo
44	31	66.0	151	13	09PUC3	09PUC3 brachydanio
45	31	66.0	151	16	09PFC3	09PFC3 xyella fas

ALIGNMENTS

RESULT 1
ID 066691 PRELIMINARY: PRT: 850 AA.

AC 066691: 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein AQ_367.
GN AQ_367.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_Taxid=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
DR EMBL: AE000686; AAC06656.1; -
DR InterPro: IPR001633; EAL.
DR InterPro: IPR001630; GDEF.
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00563; EAL. 1.
DR Pfam: PF00990; GDEF. 1.
DR SMART: SM00267; DUF1. 1.
DR SMART: SM00052; DUF2. 1.
DR TIGRFAMs: TIGR00254; GDEF. 1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN.1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 850 AA; 97669 MW; C02DCEP91DNCDB0 CRC64;

Query Match 80.9%; Score 38; DB 16; Length 850;
Best Local Similarity 54.5%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXPYEE 11
 :||:||||
 DB 227 ISREMDYPYEE 237

RESULT 2
 ID 080621 PRELIMINARY; PRT; 474 AA.
 AC 080621;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE Alpha-galactosidase.
 GN MEIA OR ATU4665 OR AGR_L436.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608550; PubMed-11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Str.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-U., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tinger S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.,
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608551; PubMed-11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Hommel K., Gordon J., Vaubin M., Iarchouk O., Papp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lapas C., Marzelz B.,
 RA Planagan C., Crowell C., Curson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009395; AAI45459.1; ALT-INIT.
 DR EMBL: AE008220; AAK8782.1; -
 KN Complete proteome.
 SQ SEQUENCE 474 AA; 53483 MW; 16C1085B5A12B21B CRC64;

Query Match 78.7%; Score 37; DB 16; Length 474;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELXPYEE 11
 ||:|||||
 DB 200 LAHDLDPYEE 210

RESULT 3
 ID 09X148 PRELIMINARY; PRT; 1451 AA.
 AC 09X148;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Pg1.15 protein.
 GN Pg1.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Ienz C., Liu S.,
 RA Lee J., Liu A., Li J., Kremetska I., Inros J., Gonzalez A.,
 RA Alafati H., Araujo R., Brookes S., Ruehler E., Chao Q., Conn L.,
 RA Conway A.B., Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C.,
 RA Rowley D., Shin P., Walker M., Davis R.W., Ecker J.R.,
 RA Federstepiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC P911 sequence.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007591; AAD39650.1; -
 DR Interpro: IPR003593; AAA_Atpase.
 DR Interpro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Prodom: PD000006; ABC_transport; 1.
 DR SMART: SM00382; AAA; 1.
 DR ATP-binding.
 SQ SEQUENCE 1451 AA; 163602 MW; 0D35414D9C370A85 CRC64;

Query Match 76.6%; Score 36; DB 10; Length 1451;
 Best Local Similarity 54.5%; Pred. No. 71;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELXPYEE 11
 ||||:|
 DB 491 LSNELSPYDK 501

RESULT 4
 ID 044755 PRELIMINARY; PRT; 634 AA.
 AC 044755;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ORP37.
 GN BB0267.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RL Dunn J.T., Butler-Loffredo L., Kieletzawa J., Medalle J., Luit B.J.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE-98065943; PubMed-9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Ishiguro R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
 RT burgdorferi.";
 RL Nature 380:580-586(1997).
 DR EMBL: U43739; AAA85591.1; -
 DR TIGR: BB0267; -
 KN Complete proteome.
 SQ SEQUENCE 634 AA; 70752 MW; E8AFB112B121F3B8 CRC64;

Query Match 74.5%; Score 35; DB 16; Length 634;
 Best Local Similarity 60.0%; Pred. No. 48;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 SXELXPYEE 11

Db 43 SLESLVPYKE 52

RESULT 5

Q9U052

PRELIMINARY;

PRT: 1028 AA.

AC Q9U052: 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 GN HNB-3.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN 11

SEQUENCE FROM N.A.

RA MEDLINE-98146168; PubMed-9486763;

RT Kamei Y., Tsutsumi O., Taketani Y., Watanabe K.;

RT "cDNA cloning and chromosomal localization of neural adhesion

RT molecule, NB-3 in human.";

RL J. Neurosci. Res. 51:275-283(1998).

DR EMBL; AB003592; BAA82612.1; .

DR HSSP; P20241; ICFB.

DR InterPro: IPR003439; ABC_transportr.

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_1Ike.

DR InterPro: IPR003006; Ig_MHC.

DR Pfam: PF00047; fn3; 4.

DR Pfam: PF00047; Ig; 6; .

DR SMART: SM00408; IGC2; 5.

DR SMART: SM00410; IG_1Ike; 1.

DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.

KW Immunoglobulin domain.

SQ SEQUENCE 1028 AA; 113956 MW; 885A2ED2F2936A6 CRC64;

Query Match

Best Local Similarity 74.5%; Score 35; DB 4; Length 1028;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SXELXIPYEE 11

Db 959 SAEILVPEEE 968

RESULT 6

Q9JMB8

PRELIMINARY;

PRT: 1028 AA.

AC Q9JMB8:

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Neural recognition molecule NB-3.

GN C2M6 OR MNB-3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN 11

SEQUENCE FROM N.A.

RA MEDLINE-20183687; PubMed-10717476;

RT Lee S., Takeda Y., Kawano H., Hosoya H., Nomoto M., Fujimoto D.,

RT Takahashi N., Watanabe K.;

RT "Expression and regulation of a gene encoding neural recognition

RT molecule NB-3 of the contactin/F3 subgroup in mouse brain.";

RL EMBL; AB032602; BAA92367.1; .

DR HSSP; P20241; ICFB.

DR MGI:1858223; Cntn6.

DR InterPro: IPR003439; ABC_transportr.

DR InterPro: IPR003961; FN_III.

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_1Ike.

DR InterPro: IPR003006; Ig_MHC.

DR Pfam: PF00047; fn3; 4.

DR Pfam: PF00047; Ig; 6.

DR SMART: SM00060; FN3; 4.

DR SMART: SM00408; IGC2; 5.

DR SMART: SM00410; IG_1Ike; 1.

DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.

DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.

KW Immunoglobulin domain.

SQ SEQUENCE 1028 AA; 113759 MW; ABB59F86CA3978F CRC64;

Query Match

Best Local Similarity 74.5%; Score 35; DB 11; Length 1028;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SXELXIPYEE 11

Db 959 SAEILVPEEE 968

RESULT 7

P97528

PRELIMINARY;

PRT: 1028 AA.

AC P97528:

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE NB-3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN 11

SEQUENCE FROM N.A.

RA STRAIN-WISTAR; TISSUE-BRAIN;

RA MEDLINE-97101230; PubMed-8945756;

RT Ogawa J., Kaneko H., Masuda T., Nagata S., Hosoya H., Watanabe K.;

RT "Novel neural adhesion molecules in the Contactin/F3 subgroup of the

RT immunoglobulin superfamily: Isolation and characterization of cDNAs

RT from rat brain.";

RL Neurosci. Lett. 218:173-176(1996).

DR EMBL; D87248; HAA18320.1; .

DR InterPro: IPR004439; ABC_transportr.

DR InterPro: IPR003961; FN_III.

DR InterPro: IPR003598; Ig_C2.

DR InterPro: IPR003600; Ig_1Ike.

DR Pfam: PF00047; fn3; 4.

DR Pfam: PF00047; Ig; 6.

DR SMART: SM00060; FN3; 4.

DR SMART: SM00408; IGC2; 5.

DR SMART: SM00410; IG_1Ike; 1.

DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.

KW Immunoglobulin domain.

SQ SEQUENCE 1028 AA; 114065 MW; 47EFD8A370CF4923 CRC64;

Query Match

Best Local Similarity 74.5%; Score 35; DB 11; Length 1028;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 SXELXIPYEE 11

Db 959 SAEILVPEEE 968

RESULT 8

Q94849

PRELIMINARY;

PRT: 274 AA.

AC Q94849;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Yolk protein 1 (Fragment).
 GN YPL.
 OS Drosophila buzzatii (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7264;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96120851; PubMed=8548285;
 RA Kambayse M.P., Ho K.F., Craddock E.M., Piano F., Parisi M.,
 RA Cohen J.;
 RT "Pattern of ecological shifts in the diversification of Hawaiian
 RL Drosophila inferred from a molecular phylogeny.";
 RL Curr. Biol. 5:1129-1139(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ho K.F.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U52126; AAC47249.1; -;
 DR FlyBase; Fg90016167; DmuzzYPL.
 DR InterPro; IPR000734; Lipase.
 DR Pfam; PF00151; Lipase; 1.
 FT NON_TER 1
 FT 274
 SQ SEQUENCE 274 AA; 30637 MW; 772CFE9AE771BA8 CRC64;
 Query Match 72.3%; Score 34; DB 5; Length 274;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSXELIPYE 10
 DB 220 LTSELEPYE 229

RESULT 9
 Q9RB89 PRELIMINARY; PRT; 291 AA.
 AC Q9RB89;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate
 DE lyase.
 GN LIVE OR TTE0933.
 OS Thermotogaobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermotogaobacteriales; Thermotogaobacteriaceae; Thermotogaobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL; AE013059; AAM24189.1; -;
 KW Transferase; Aminotransferase; Complete proteome.
 SQ SEQUENCE 291 AA; 32346 MW; 105072D27AB06FFE CRC64;
 Query Match 72.3%; Score 34; DB 16; Length 291;
 Best Local Similarity 54.5%; Pred. No. 35;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSXELIPYE 11
 DB 220 LTSELEPYE 229

DB 218 IARELNIPFE 228
 RESULT 10
 Q9YKM5 PRELIMINARY; PRT; 381 AA.
 AC Q9YKM5;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Ldorf-102 peptide.
 OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124785; PubMed=9887315;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J.M., Rohrmann G.F.;
 RT "Sequence and analysis of the genome of a baculovirus pathogenic for
 RT Lymantria dispar";
 RL Virology 253:17-34(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF081810; AAC70288.1; -;
 SQ SEQUENCE 381 AA; 43418 MW; EA63951C43C3ACA6 CRC64;
 Query Match 72.3%; Score 34; DB 12; Length 381;
 Best Local Similarity 75.0%; Pred. No. 46;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 ELXIPYE 11
 DB 41 ELDPYE 48

RESULT 11
 Q9XWY5 PRELIMINARY; PRT; 738 AA.
 AC Q9XWY5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Y66A7A.5 protein.
 GN Y66A7A.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shomkhen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).
 RN [3]

RP SEQUENCE FROM N.A.
 RA Sulston J.E.;
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
 RP
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851915;
 RA none;
 RT "genome sequence of the nematode C.elegans: A platform for
 investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL: AL032622; CAA21499.1; -;
 DR EMBL: AL590342; CAC35859.1; -;
 DR InterPro: IPR001356; Homeobox.
 DR SMART: SM00389; HOX; 2.
 DR PROSITE: PSS0071; HOMEBOX.2; 2.
 SQ SEQUENCE 738 AA; 87325 MW; FFA0174853A0E2A7 CRC64;

Query Match 72.3%; Score 34; DB 5; Length 738;
 Best Local Similarity 60.0%; Pred. No. 93;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELXPYEE 10
 Db 673 LAFEDLPYEE 682

RESULT 12
 Q92FP6 PRELIMINARY; PRT; 233 AA.

AC Q92FP6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein lin0057.

GN LIN0057.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed-11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Cheouani F., Couve E., de Daruvar A., Deboux P.,
 Doman E., Dominguez-Bernal G., Duchaud E., Durant L., Dussaugel O.,
 Enliat K.-D., Esili H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haul J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kutapkat G.,
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstrek C., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 Vaquer-Boland J.-A., Voss H., Weiland J., Cossart P.;
 "Comparative genomics of Listeria species";
 RT Science 294:849-852(2001).
 DR EMBL: AL596163; CAC95290.1; -;
 DR Listlist: LIN0057; -;
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 233 AA; 27625 MW; E1E405DC36048EB CRC64;

Query Match 70.2%; Score 33; DB 16; Length 233;
 Best Local Similarity 63.6%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 LSXELXPYEE 11
 Db 221 LFELEFPYEE 231

RESULT 13
 OBYA03 PRELIMINARY; PRT; 233 AA.

AC OBYA03;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein lin0064.
 GN LMO0064.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-BGD-B / SEROVAR 1/2A;
 RX MEDLINE-21537279; PubMed-11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Cheouani F., Couve E., de Daruvar A., Deboux P.,
 Doman E., Dominguez-Bernal G., Duchaud E., Durant L., Dussaugel O.,
 Enliat K.-D., Esili H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haul J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kutapkat G.,
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstrek C., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 Vaquer-Boland J.-A., Voss H., Weiland J., Cossart P.;
 "Comparative genomics of Listeria species";
 RT Science 294:849-852(2001).
 DR EMBL: AL591973; CAC98279.1; -;
 DR Listlist: LMO0064; -;
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 233 AA; 27647 MW; 2ED2B06E4FB2C481 CRC64;

Query Match 70.2%; Score 33; DB 16; Length 233;
 Best Local Similarity 63.6%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 LSXELXPYEE 11
 Db 221 LFELEFPYEE 231

RESULT 14
 Q989F3 PRELIMINARY; PRT; 461 AA.

AC Q989F3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alpha-dactosinase.

GN MR6450.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-MAF30309;
 RX MEDLINE-21082930; PubMed-11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu K., Kato T., Sasamoto S.,
 Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
 Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
 Mochizuki Y., Nakayama S., Nakazaki N., Shingo S., Sugimoto M.,
 Tatemuchi C., Yamada M., Tabata S.;
 "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti";
 RT DNA Res. 7:331-338(2000).

DR EMBL: AP003009; BAB52744.1; -;
 DR InterPro: IPR01088; GH_4.
 DR Pfam: PF02056; Glyco_hydro.4; 1.
 DR PRINTS: PR00732; GHYDRASE4.
 DR PROSITE: P006892; GH_4; 1.

DR PROSITE: PSS1324; GLYCOSYL_HYDROL_F4; UNKNOWN.1.
 KW Complete proteome.

SQ SEQUENCE 461 AA; 51818 MW; 6C29F18157E4F6EC CRC64;

Query Match 70.2%; Score 33; DB 16; Length 461;

Best Local Similarity 45.5%; Pred. No. 92;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11

1:1:1:1:1

Db 188 LAHDLSTPYDE 198

RESULT 15

Q9Q4W5

ID 09Q4W5 PRELIMINARY; PRT; 465 AA.

AC 09Q4W5;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Polyprotein (Fragment).

OS Ornithogalum mosaic virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;

OC Potyvirus

OX NCBI_TaxID-12204;

RN [1]

RP SEQUENCE FROM N.A.

RA Mackenzie A., Gibbs A.;

RT "Potyviruses in Australia."

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF203528; AAF22752.1; -;

DR InterPro; IPR001592; Poly_coat.

DR InterPro; IPR001205; RNA_pol_P3D.

DR Pfam; PF00767; Poly_coat; 1.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

FT NON_TER

FT 1

SQ SEQUENCE 465 AA; 53264 MW; 3D4B484C259F636 CRC64;

Query Match

Best Local Similarity

Matches 6; Conservative

QY 4 ELXIPYEE 11

1:1:1:1:1

Db 27 KLSIPYEE 34

Score 33; DB 12; Length 465;

Pred. No. 93;

Mismatches 1; Indels 0; Gaps 0;

Search completed: January 15, 2003, 12:35:38

Job time : 35.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 16, 2003, 08:03:43 ; Search time 1525.14 Seconds
(without alignments)
209.902 Million cell updates/sec

Title: US-09-823-649a-2
Perfect score: 47
Sequence: 1 LSXELXIPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -O -f -s -t -m -n -p -r -u -v -w -x -y -z -A -B -C -D -E -F -G -H -I -J -K -L -M -N -O -P -Q -R -S -T -U -V -W -X -Y -Z -AA -AB -AC -AD -AE -AF -AG -AH -AI -AJ -AK -AL -AM -AN -AO -AP -AQ -AR -AS -AT -AU -AV -AW -AX -AY -AZ -BA -BB -BC -BD -BE -BF -BG -BH -BI -BJ -BK -BL -BM -BN -BO -BP -BQ -BR -BS -BT -BU -BV -BW -BX -BY -BZ -CA -CB -CC -CD -CE -CF -CG -CH -CI -CJ -CK -CL -CM -CN -CO -CP -CQ -CR -CS -CT -CU -CV -CW -CX -CY -CZ -DA -DB -DC -DD -DE -DF -DG -DH -DI -DJ -DK -DL -DM -DN -DO -DP -DQ -DR -DS -DT -DU -DV -DW -DX -DY -DZ -EA -EB -EC -ED -EE -EF -EG -EH -EI -EJ -EK -EL -EM -EN -EO -EP -EQ -ER -ES -ET -EU -EV -EW -EX -EY -EZ -FA -FB -FC -FD -FE -FF -FG -FH -FI -FJ -FK -FL -FM -FN -FO -FP -FQ -FR -FS -FT -FU -FV -FW -FX -FY -FZ -GA -GB -GC -GD -GE -GF -GG -GH -GI -GJ -GK -GL -GM -GN -GO -GP -GQ -GR -GS -GT -GU -GV -GW -GX -GY -GZ -HA -HB -HC -HD -HE -HF -HG -HH -HI -HJ -HK -HL -HM -HN -HO -HP -HQ -HR -HS -HT -HU -HV -HW -HX -HY -HZ -IA -IB -IC -ID -IE -IF -IG -IH -II -IJ -IK -IL -IM -IN -IO -IP -IQ -IR -IS -IT -IU -IV -IW -IX -IY -IZ -JA -JB -JC -JD -JE -JF -JG -JH -JI -JJ -JK -JL -JM -JN -JO -JP -JQ -JR -JS -JT -JU -JV -JW -JX -JY -JZ -KA -KB -KC -KD -KE -KF -KG -KH -KI -KJ -KK -KL -KM -KN -KO -KP -KQ -KR -KS -KT -KU -KV -KW -KX -KY -KZ -LA -LB -LC -LD -LE -LF -LG -LH -LI -LJ -LK -LL -LM -LN -LO -LP -LQ -LR -LS -LT -LU -LV -LW -LX -LY -LZ -MA -MB -MC -MD -ME -MF -MG -MH -MI -MJ -MK -ML -MM -MN -MO -MP -MQ -MR -MS -MT -MU -MV -MW -MX -MY -MZ -NA -NB -NC -ND -NE -NF -NG -NH -NI -NJ -NK -NL -NM -NO -NP -NQ -NR -NS -NT -NU -NV -NW -NX -NY -NZ -OA -OB -OC -OD -OE -OF -OG -OH -OI -OJ -OK -OL -OM -ON -OO -OP -OQ -OR -OS -OT -OU -OV -OW -OX -OY -OZ -PA -PB -PC -PD -PE -PF -PG -PH -PI -PJ -PK -PL -PM -PN -PO -PP -PQ -PR -PS -PT -PU -PV -PW -PX -PY -PZ -QA -QB -QC -QD -QE -QF -QG -QH -QI -QJ -QK -QL -QM -QN -QO -QP -QQ -QR -QS -QT -QU -QV -QW -QX -QY -QZ -RA -RB -RC -RD -RE -RF -RG -RH -RI -RJ -RK -RL -RM -RN -RO -RP -RQ -RR -RS -RT -RU -RV -RW -RX -RY -RZ -SA -SB -SC -SD -SE -SF -SG -SH -SI -SJ -SK -SL -SM -SN -SO -SP -SQ -SR -SS -ST -SU -SV -SW -SX -SY -SZ -TA -TB -TC -TD -TE -TF -TG -TH -TI -TJ -TK -TL -TM -TN -TO -TP -TQ -TR -TS -TT -TU -TV -TW -TX -TY -TZ -UA -UB -UC -UD -UE -UF -UG -UH -UI -UJ -UK -UL -UM -UN -UO -UP -UQ -UR -US -UT -UU -UV -UW -UX -UY -UZ -VA -VB -VC -VD -VE -VF -VG -VH -VI -VJ -VK -VL -VM -VN -VO -VP -VQ -VR -VS -VT -VU -VV -VW -VX -VY -VZ -WA -WB -WC -WD -WE -WF -WG -WH -WI -WJ -WK -WL -WM -WN -WO -WP -WQ -WR -WS -WT -WU -WV -WW -WX -WY -WZ -XA -XB -XC -XD -XE -XF -XG -XH -XI -XJ -XK -XL -XM -XN -XO -XP -XQ -XR -XS -XT -XU -XV -XW -XX -XY -XZ -YA -YB -YC -YD -YE -YF -YG -YH -YI -YJ -YK -YL -YM -YN -YO -YP -YQ -YR -YS -YT -YU -YV -YW -YX -YY -YZ -ZA -ZB -ZC -ZD -ZE -ZF -ZG -ZH -ZI -ZJ -ZK -ZL -ZM -ZN -ZO -ZP -ZQ -ZR -ZS -ZT -ZU -ZV -ZW -ZX -ZY -ZZ

Database : GenEmbl:1*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	44	93.6	1079	8	AF315118	AP315118 Populus e
2	44	93.6	1600	6	AR023967	AR023967 Sequence
3	44	93.6	1600	6	AR054704	AR054704 Sequence
4	44	93.6	1600	6	AR061700	AR061700 Sequence
5	44	93.6	1600	6	AR061941	AR061941 Sequence
6	44	93.6	1600	6	AR063642	AR063642 Sequence
7	44	93.6	1600	6	AR086123	AR086123 Sequence
8	44	93.6	1600	6	AR089553	AR089553 Sequence
9	44	93.6	1600	6	AR093638	AR093638 Sequence
10	44	93.6	1600	6	AR193577	AR193577 Sequence
11	44	93.6	1600	6	AR206175	AR206175 Sequence
12	44	93.6	1600	6	IZ4097	IZ4097 Sequence 12
13	44	93.6	1600	6	I38585	I38585 Sequence 12
14	44	93.6	1600	6	I76140	I76140 Sequence 12
15	44	93.6	1682	6	AR112315	AR112315 Sequence
16	44	93.6	1696	6	AR112316	AR112316 Sequence
17	44	93.6	1812	6	AX093143	AX093143 Sequence
18	44	93.6	1812	6	AX093145	AX093145 Sequence
19	44	93.6	2433	6	AX317584	AX317584 Sequence
20	44	93.6	2445	6	AX317588	AX317588 Sequence
21	44	93.6	2445	6	AX317572	AX317572 Sequence
22	44	93.6	2445	6	AX317576	AX317576 Sequence
23	44	93.6	2445	6	AX317578	AX317578 Sequence
24	44	93.6	2480	6	AX404968	AX404968 Sequence
25	44	93.6	2483	6	AX404966	AX404966 Sequence
26	44	93.6	2484	6	AX404972	AX404972 Sequence
27	44	93.6	2493	6	AX023936	AX023936 Sequence
28	44	93.6	2493	6	AX317586	AX317586 Sequence
29	44	93.6	2493	6	I15437	I15437 Sequence 5
30	44	93.6	2496	6	AR023961	AR023961 Sequence
31	44	93.6	2496	6	AR054698	AR054698 Sequence
32	44	93.6	2496	6	AR061694	AR061694 Sequence
33	44	93.6	2496	6	AR061935	AR061935 Sequence
34	44	93.6	2496	6	AR063636	AR063636 Sequence
35	44	93.6	2496	6	AR086117	AR086117 Sequence
36	44	93.6	2496	6	AR089547	AR089547 Sequence
37	44	93.6	2496	6	AR093632	AR093632 Sequence
38	44	93.6	2496	6	AR193571	AR193571 Sequence
39	44	93.6	2496	6	AR206169	AR206169 Sequence
40	44	93.6	2496	6	AX317151	AX317151 Sequence
41	44	93.6	2496	6	AX317344	AX317344 Sequence
42	44	93.6	2496	6	AX317349	AX317349 Sequence
43	44	93.6	2496	6	AX317361	AX317361 Sequence
44	44	93.6	2496	6	AX317365	AX317365 Sequence
45	44	93.6	2496	6	I24091	I24091 Sequence 2

ALIGNMENTS

RESULT 1

AF315118/c
 LOCUS AF315118 1079 bp mRNA linear PLN 20-FEB-2001
 DEFINITION Populus euphratica clone PSRI salt-induced mRNA sequence.
 ACCESSION AF315118
 VERSION AF315118.1 GI:12964697
 KEYWORDS
 SOURCE
 ORGANISM
 Populus euphratica.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids 1; Malpighiales; Salicaceae; Populus.
 REFERENCE
 AUTHORS Huiming, Z. and Xin, S.
 TITLE Salt inducible gene in Populus euphratica
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 1079)
 AUTHORS Huiming, Z. and Xin, S.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-2000) Biology, Beijing Forestry University,
 Tsinghua East Rd. No.35, Beijing 100083, China
 FEATURES
 source
 1..1079
 /organism="Populus euphratica"
 /db_xref="taxon:75702"
 /clone="PSRI"
 /tissue_type="root tips"
 /dev_stage="two years old"
 /note="Isolated using suppression subtractive
 hybridization of uninduced vs. salt-induced root tips;
 cloned using RACE"
 1..1079
 misc_feature
 /note="salt-induced transcript"
 BASF COUNT 158 a 355 c 369 g 197 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.31 Length: 1079
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 8 Gaps: 0
 US-09-823-649A-2 (1-11) x AF315118 (1-1079)
 QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
 ||||| ||||| ||||| ||||| |||||
 Db 468 CTCTCCAGAGAGCTAGCCATCCCTTACGAGAG 436
 RESULT 2
 LOCUS AR023967 1600 bp DNA linear PAT 05-DEC-1998
 DEFINITION Sequence 12 from patent US 5795763.
 ACCESSION AR023967
 VERSION AR023967.1 GI:3977261
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1600)
 AUTHORS Dahlberg, J.E., Lyamichev, V.I. and Brow, M. Ann.D.
 TITLE Synthesis-deficient thermostable DNA polymerase
 JOURNAL Patent: US 5795763-A 12 18-AUG-1998;
 FEATURES
 source
 1..1600
 /organism="unknown"
 BASF COUNT 286 a 519 c 553 g 242 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.84 Length: 1600
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 Percent Similarity: 81.82% Gaps: 0

Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 6 Gaps: 0
 US-09-823-649A-2 (1-11) x AR023967 (1-1600)
 QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
 ||||| ||||| ||||| ||||| |||||
 Db 1133 CTCTCCAGAGAGCTAGCCATCCCTTACGAGAG 1165
 RESULT 3
 LOCUS AR054704 1600 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 12 from patent US 5837450.
 ACCESSION AR054704
 VERSION AR054704.1 GI:5980281
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1600)
 AUTHORS Dahlberg, J.E., Lyamichev, V.I. and Brow, M. Ann.D.
 TITLE Detection of target nucleic acid molecules using thermostable 5'
 nuclease
 JOURNAL Patent: US 5837450-A 12 17-NOV-1998;
 FEATURES
 source
 1..1600
 /organism="unknown"
 BASF COUNT 286 a 519 c 553 g 242 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.84 Length: 1600
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 6 Gaps: 0
 US-09-823-649A-2 (1-11) x AR054704 (1-1600)
 QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
 ||||| ||||| ||||| ||||| |||||
 Db 1133 CTCTCCAGAGAGCTAGCCATCCCTTACGAGAG 1165
 RESULT 4
 LOCUS AR061700 1600 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 12 from patent US 5843654.
 ACCESSION AR061700
 VERSION AR061700.1 GI:5989391
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1600)
 AUTHORS Heisler, L.M., Fors, L. and Brow, M. Ann.D.
 TITLE Rapid detection of mutations in the p53 gene
 JOURNAL Patent: US 5843654-A 12 01-DEC-1998;
 FEATURES
 source
 1..1600
 /organism="unknown"
 BASF COUNT 286 a 519 c 553 g 242 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.84 Length: 1600
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 6 Gaps: 0

US-09-823-649a-2 (1-11) x AR061700 (1-1600)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11

Db 1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165

RESULT 5

AR061941

LOCUS AR061941 1600 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 12 from patent US 5843669.

ACCESSION AR061941

VERSION AR061941.1 GI:5989632

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1600)

AUTHORS

Kaiser, M.W., Lyamichev, V.I. and Lyamichev, N.

Cleavage of nucleic acid using thermostable methanococcus

janaschii PEN-1 endonucleases

Patent: US 5843669-A 12 01-DEC-1998;

FEATURES

source

ORIGIN

BASE COUNT

286 a

519 c

553 g

242 t

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-823-649a-2 (1-11) x AR061941 (1-1600)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11

Db 1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165

RESULT 6

AR063642

LOCUS AR063642

DEFINITION Sequence 12 from patent US 5846717.

ACCESSION AR063642

VERSION AR063642.1 GI:5992950

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1600)

AUTHORS

Brow, M. Ann. D., Hall, J. Steven, Grotelueschen, J., Lyamichev, V.,

Oliver, D. Michael, and Prudent, J. Robert.

Detection of nucleic acid sequences by invader-directed cleavage

Patent: US 5846717-A 12 08-DEC-1998;

FEATURES

source

ORIGIN

BASE COUNT

286 a

519 c

553 g

242 t

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-823-649a-2 (1-11) x AR063642 (1-1600)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11

Db 1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165

RESULT 7

AR086123

LOCUS AR086123

DEFINITION Sequence 12 from patent US 5985557.

ACCESSION AR086123

VERSION AR086123.1 GI:10012889

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1600)

AUTHORS

Prudent, J.R., Hall, J.G., Lyamichev, V.I., Brow, M. Ann. D. and

Dahlberg, J.E.

Invasive cleavage of nucleic acids

Patent: US 5985557-A 12 16-NOV-1999;

FEATURES

source

ORIGIN

BASE COUNT

286 a

519 c

553 g

242 t

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-823-649a-2 (1-11) x AR086123 (1-1600)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11

Db 1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165

RESULT 8

AR089553

LOCUS AR089553

DEFINITION Sequence 12 from patent US 5994069.

ACCESSION AR089553

VERSION AR089553.1 GI:10016310

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1600)

AUTHORS

Hall, J.G., Lyamichev, V.I., Past, A. L., and Brow, M. Ann. D.

Detection of nucleic acids by multiple sequential invasive

cleavages

Patent: US 5994069-A 12 30-NOV-1999;

FEATURES

source

ORIGIN

BASE COUNT

286 a

519 c

553 g

242 t

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-823-649a-2 (1-11) x AR089553 (1-1600)

QY 1 Leuser***GluLeu***IleProTyrGluGlu 11

Db 1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165

RESULT 9
AR093638
LOCUS AR093638 1600 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 12 from patent US 6001567.
ACCESSION AR093638
VERSION AR093638.1 GI:10020387
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Brow,M,Ann.D., Hall,J.G., Lyamichev,V.I.,
Olive,D,Michael, and Prudent,J,Robert.
TITLE Detection of nucleic acid sequences by invader-directed cleavage
JOURNAL Patent: US 6001567-A 12 14-DEC-1999;
FEATURES
Location/Qualifiers
source 1..1600
BASE COUNT 286 a 519 c 553 g 242 t
ORIGIN
Alignment Scores:
Pred. No.: 1.84 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-2 (1-11) x AR093638 (1-1600)
QY 1 LeuSer**GlUleu**IlleProTyrglUgu 11
Db 1133 CTCTCCAGAGACTAGCATCCCTTACGAGAG 1165
RESULT 10
AR193577
LOCUS AR193577 1600 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 12 from patent US 6348314.
ACCESSION AR193577
VERSION AR193577.1 GI:20240169
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Prudent,J,R., Hall,J.G., Lyamichev,V.I., Brow,M,Ann.D. and
Dahiberg,J,E.
TITLE Invasive cleavage of nucleic acids
JOURNAL Patent: US 6348314-A 12 19-APR-2002;
FEATURES
Location/Qualifiers
source 1..1600
BASE COUNT 286 a 519 c 553 g 242 t
ORIGIN
Alignment Scores:
Pred. No.: 1.84 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-2 (1-11) x AR193577 (1-1600)
QY 1 LeuSer**GlUleu**IlleProTyrglUgu 11
Db 1133 CTCTCCAGAGACTAGCATCCCTTACGAGAG 1165
RESULT 11
AR206175
LOCUS AR206175 1600 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 12 from patent US 6372424.
ACCESSION AR206175
VERSION AR206175.1 GI:21504699
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Brow,M,Ann.D., Lyamichev,V.I. and Olive,D,Michael.
TITLE Rapid detection and identification of pathogens
JOURNAL Patent: US 6372424-A 12 16-APR-2002;
FEATURES
Location/Qualifiers
source 1..1600
BASE COUNT 286 a 519 c 553 g 242 t
ORIGIN
Alignment Scores:
Pred. No.: 1.84 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-2 (1-11) x AR206175 (1-1600)
QY 1 LeuSer**GlUleu**IlleProTyrglUgu 11
Db 1133 CTCTCCAGAGACTAGCATCCCTTACGAGAG 1165
RESULT 12
I24097
LOCUS I24097 1600 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 12 from patent US 5541311.
ACCESSION I24097
VERSION I24097.1 GI:1603967
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Dahiberg,J,E., Lyamichev,V.I. and Brow,M,Ann.D.
TITLE Nucleic acid encoding synthesis-deficient thermostable DNA
polymerase
JOURNAL Patent: US 5541311-A 12 30-JUL-1996;
FEATURES
Location/Qualifiers
source 1..1600
BASE COUNT 286 a 519 c 553 g 242 t
ORIGIN
Alignment Scores:
Pred. No.: 1.84 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-2 (1-11) x I24097 (1-1600)
QY 1 LeuSer**GlUleu**IlleProTyrglUgu 11
Db 1133 CTCTCCAGAGACTAGCATCCCTTACGAGAG 1165
RESULT 13
I38585
LOCUS I38585 1600 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 12 from patent US 5614402.
ACCESSION I38585
VERSION I38585.1 GI:2084639
KEYWORDS

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 11:21:44 ; Search time 130.143 Seconds
(without alignments)
190.345 Million cell updates/sec

Title: US-09-823-649A-2
Perfect score: 47
Sequence: 1 LSXELXPYEE 11

Scoring table:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -v -
-MODEL=frame_p2n.model -DEV=rxh
-O=/cgn1/USPTO.spool/US09823649/runcat.14012003.151001.29119/epq.query.fasta.1.1393
-DB=N_Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blastsum2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.qcgn.1.1.0.0runat.14012003.151001.29119 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	44	93.6	1600	16	AA080745	5' Nuclease from T
2	44	93.6	1600	17	AA027686	Mutant Thermus aqu
3	44	93.6	1600	18	AA070347	Synthesis deficient
4	44	93.6	1600	18	AA076647	Tag gene 5' nuclea
5	44	93.6	1600	19	AA05786	Thermus aquaticus
6	44	93.6	1600	19	AAV53855	Nucleotide sequenc
7	44	93.6	1600	20	AAV63407	DNA sequence of a
8	44	93.6	1635	13	AA024010	Mutant thermostabi
9	44	93.6	1635	13	AA023998	Mutant thermostabi
10	44	93.6	1635	13	AA024322	Mutant thermostabi
11	44	93.6	1635	13	AA024328	Mutant thermostabi
12	44	93.6	1682	21	AA029460	Thermotable DNA d
13	44	93.6	1686	18	AA047959	Mutant Tag polymer
14	44	93.6	1686	18	AA047961	Mutant Tag polymer
15	44	93.6	1686	20	AA027130	FY2 polymerase cod
16	44	93.6	1686	20	AA027132	FY4 polymerase cod
17	44	93.6	1689	18	AA047860	Mutant Tag polymer
18	44	93.6	1689	20	AA027131	FY3 polymerase cod
19	44	93.6	1696	21	AA029461	Thermotable DNA d
20	44	93.6	1794	17	AA032327	Thermus flavus DNA
21	44	93.6	1794	17	AA027255	DNA polymerase I e
22	44	93.6	1812	22	AA000717	Chimeric CaulTag DN
23	44	93.6	1812	22	AA000718	Chimeric CaulTag DN
24	44	93.6	1896	13	AA023997	Mutant thermostabi
25	44	93.6	1899	13	AA024009	Mutant thermostabi
26	44	93.6	1899	13	AA024321	Mutant thermostabi
27	44	93.6	1899	13	AA024327	Mutant thermostabi
28	44	93.6	1899	24	AA024819	Sulfolobus solfata
29	44	93.6	1904	24	AA024822	Sulfolobus solfata
30	44	93.6	1965	24	AA024823	Polylysine (PL) -
31	44	93.6	2043	13	AA024008	Mutant thermostabi
32	44	93.6	2043	13	AA023996	Mutant thermostabi
33	44	93.6	2043	13	AA024320	Mutant thermostabi
34	44	93.6	2043	13	AA024326	Mutant thermostabi
35	44	93.6	2274	13	AA023894	Mutant thermostabi
36	44	93.6	2277	13	AA024007	Mutant thermostabi
37	44	93.6	2277	13	AA024007	Mutant thermostabi
38	44	93.6	2277	13	AA024325	Mutant thermostabi
39	44	93.6	2364	13	AA023995	Mutant thermostabi
40	44	93.6	2367	13	AA024006	Mutant thermostabi
41	44	93.6	2370	13	AA024012	Mutant thermostabi
42	44	93.6	2370	13	AA024324	Mutant thermostabi
43	44	93.6	2493	13	AA024005	Mutant thermostabi
44	44	93.6	2496	16	AA080750	T. flavus DNA-poly
45	44	93.6	2496	19	AA065780	Thermus flavus DNA

ALIGNMENTS

RESULT 1
AA080745
ID AA080745 standard; DNA; 1600 Bp.
AC AA080745:
XX
XX 19-JUL-1995 (first entry)
DE 5' Nuclease from Tag DNAP.
XX
XX DNA-polymerase; DNAP; Tag; DNA cleavage; RNA cleavage; 5' nuclease; ds.
OS Thermus aquaticus YR-1.
XX
XX WO9429482-A.
XX
XX PD 22-DEC-1994.
XX

```

PF 06-JUN-1994; 94WO-US06253.
XX
PR 04-JUN-1993; 93US-0073384.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Lyamichev VI;
XX
DR WPI; 1995-036504/05.
XX
PT 5' Nuclease(s) derived from thermostable DNA polymerase(s) - have
PT cleavage activity with reduced synthetic ability, used for
PT detection of specific target sequences.
XX
PS Claim 5; Page 95-96; 159pp; English.
XX
CC The Taq DNA-polymerase (DNAP) gene was amplified by PCR. Amplified
CC fragments were ligated into pT7018 vector, which contains the
CC hybrid trp-lac (lac) promoter. An amplification/selection
CC protocol was used to isolate clone 4B containing a mutated Taq DNAP
CC gene (mutTaq) (sequence given in AA080746) having normal 5' nuclease
CC activity but less than 1% of the wt Taq DNAP activity. mutTaq was
CC cut from pT7018 by EcoRI-SalI digestion and cloned into pET-3c.
CC Thus clone was digested with BstXI and BamHI. The DNA was treated
CC with DNAPcI klenow fragment and dNTPs, blunt-ended and religated,
CC resulting in an in-frame deletion of 903 nucleotides. The DNA
CC sequence of the resulting 5' nuclease is given in AA080745.
XX
SQ Sequence 1600 BP; 287 A; 518 C; 553 G; 242 T; 0 other;

Alignment Scores:
Pred. No.: 1.89 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 16 Gaps: 0

US-09-823-649A-2 (1-11) x AA080745 (1-1600)
QY 1 LeuSer***GluLeu***IleProTyrGluIn 11
Db 1133 CTCCTCCAGAGCTAGCCATCCTTAGCGAGG 1165

RESULT 2
AAT27686
ID AAT27686 standard; DNA; 1600 BP.
XX
AC AAT27686;
XX
DT 18-NOV-1996 (first entry)
XX
DE Mutant Thermus aquaticus DNA polymerase coding sequence (Clone 4F).
XX
KM p53; mutant; mutation: cleavage; nuclease; cleavage; Thermus;
KM Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
KM Staphylococcus; identification; detection; ds.
XX
OS Thermus aquaticus.
XX
FH Key Location/Qualifiers
FT 14..1600
FT CDS /*tag= a
FT /product= Mutant DNA polymerase.
XX
XX MO9615267-A1.
XX
PD 23-MAY-1996.
XX
PF 09-NOV-1995; 95WO-US14673.
XX
PR 30-AUG-1995; 95US-0520946.
PR 09-NOV-1994; 94US-0337164.

```

```

PR 09-MAR-1995; 95US-0402601.
PR 07-JUN-1995; 95US-0484956.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Fors L, Heisler LM, Lyamichev VI;
PI Oldenburg MC, Olive DM;
XX
DR WPI; 1996-259862/26.
DR P-PSDB; AAR96267.
XX
PT Cleavage of nucleic acids to detect mutation(s) - allows detection
PT esp. in human p53 gene, to identify strains of microorganisms and
PT viruses
XX
PS Example 2A; Page 257-258; 433pp; English.
XX
CC Cleavage of nucleic acids using an enzyme, especially a nuclease
CC selected from the group consisting of Cleavage (RTM) BN enzyme,
CC Thermus aquaticus DNA polymerase, Thermus thermophilus DNA
CC polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
CC Rad1/Rad10 complex. The nucleic acid substrate is preferably an
CC oligonucleotide containing a human p53 gene sequence or
CC alternatively, microbial gene sequences. Cleavage products are
CC compared to the cleavage products of reference gene sequences. The
CC method is used for detecting mutation in the human p53 gene; for
CC identifying strains of microorganisms, especially bacteria selected
CC from the group of members of the genera Campylobacter,
CC Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
CC The method may also be used for the identification of viruses,
CC especially hepatitis C virus and simian immunodeficiency virus.
CC Thermus aquaticus (Taq) DNA polymerase was amplified using two
CC primers (AAT27679, AAT27680). The Taq polymerase DNA was inserted into
CC the BamHI restriction site of the expression vector pET-3c and mutant
CC genes were created from that construct. This mutant was created
CC after the vector was digested with BstXI and BamHI. The DNA was
CC then treated with the Klenow fragment of DNAPcI to trim both 3'
CC overhangs to blunt ends which were then ligated together, resulting
CC in an in frame deletion of 903 nucleotides. The resulting mutant
CC Taq polymerase is also referred to as the Cleavage BB enzyme.
XX
SQ Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;

Alignment Scores:
Pred. No.: 1.89 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 17 Gaps: 0

US-09-823-649A-2 (1-11) x AAT27686 (1-1600)
QY 1 LeuSer***GluLeu***IleProTyrGluIn 11
Db 1133 CTCCTCCAGAGCTAGCCATCCTTAGCGAGG 1165

RESULT 3
AAT70347
ID AAT70347 standard; DNA; 1600 BP.
XX
AC AAT70347;
XX
DT 03-APR-1998 (first entry)
XX
DE Synthesis deficient Taq DNA polymerase gene clone 4F.
XX
KM DNA polymerase; taq polymerase gene; DNAP; 5' nuclease activity;
KM cleavage BB; DNA cleavage; reduced synthetic activity; ss.
XX
OS Synthetic.
XX
OS Thermus aquaticus.

```

PN US5614402-A.
XX
PD 25-MAR-1997.
XX
PF 06-JUN-1994; 94US-0254359.
XX
PR 07-DEC-1992; 92US-0986330.
PR 04-JUN-1993; 93US-0073384.
XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI Brow MAD, Dahlberg JE, Lyamichev VI;
XX WPI; 1997-201481/18.
XX
XX
XX Thermosable 5' nuclease derived from thermosable polymerase - has
PT reduced synthetic activity useful in nucleic acid detection assays
PS Example 2A; Columns 79-82; 93pp; English.
XX
XX The present sequence, clone 4F, encodes a Tag DNA polymerase enzyme
CC that has normal 5' nuclease activity, but reduced synthetic ability.
CC This sequence was derived from the muttag construct described in
CC AAT70343. The entire muttag gene was cut from the plasmid, and cloned
CC into PET-3c. This clone was digested with BstXI and BamHI, at unique
CC sites. The 3' overhang of BstXI was trimmed to a blunt end, while the 5'
CC overhang of BamHI was filled in. The blunt ends were ligated together.
CC This resulted in an in-frame deletion of 903 nucleotides. The enzyme
CC encoded by the present protein is also referred to as Cleavase BB.
XX
SQ Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.89 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 18 Gaps: 0
US-09-823-649a-2 (1-11) x AAT70347 (1-1600)
OY 1 Leuser**Glueu**IleProTyrgluGlu 11
Db 1133 CTCTCCAGAGACTAGCATCCCTTAGAGAGAG 1165
RESULT 4
AAT76647
ID AAT76647 standard; DNA; 1600 BP.
XX
AC AAT76647;
XX
XX 14-APR-1998 (first entry)
DT
XX
XX Tag gene 5' nuclease clone 3F (Cleavase BB).
DE
XX
XX Nucleic acid cleavage; DNA cleavage; RNA cleavage; 5' nuclease;
KW Tag; DNA polymerase; Cleavase BB; ds.
XX
XX Thermus aquaticus YT-1.
OS Synthetic.
OS
XX
XX WO9727214-A1.
PN
XX
XX 31-JUL-1997.
PD
XX
XX 22-JAN-1997; 97WO-US01072.
PF
XX
XX 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-0599491.
PR 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
PR 02-DEC-1996; 96US-0758314.

XX
PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX
XX Brow MAD, Dahlberg JE, Hall JG, Kaiser MW, Lyamichev VI;
PI Olive DM, Prudent JR;
XX WPI; 1997-393613/36.
XX
XX
XX Thermosable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
PT acid sequences
PS Example 2; Page 245; 457pp; English.
XX
XX This DNA sequence, denoted clone 3F or Cleavase BB, comprises an
CC altered Thermus aquaticus DNA polymerase (tag) gene in which
CC nucleotides 875-1778 of the wild-type gene coding sequence are
CC deleted. Mutant gene muttag (see AAT76643) was used as the starting
CC material for the construct. Cleavase BB is a thermosable
CC structure-specific nuclease preferred for use in nucleic acid
CC cleavage methods of the invention. Mutant genes (AAT76644-47)
CC were constructed in order to determine which portions of the Tag
CC polymerase domain can be altered without eliminating 5' nuclease
CC activity. The invention relates to means for the detection and
CC characterisation of nucleic acid (NA) sequences and variations in
CC NA. It relates to methods for forming a NA cleavage structure on
CC a target sequence and cleaving the NA cleavage structure in a
CC site-specific manner. The 5' nuclease activity of various enzymes
CC (see AAT24210-13) is used to cleave the target-dependent cleavage
CC structure, thereby indicating the presence of specific NA sequences
CC or specific variations of them.
XX
SQ Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.89 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 18 Gaps: 0
US-09-823-649a-2 (1-11) x AAT76647 (1-1600)
OY 1 Leuser**Glueu**IleProTyrgluGlu 11
Db 1133 CTCTCCAGAGACTAGCATCCCTTAGAGAGAG 1165
RESULT 5
AAV65786
ID AAV65786 standard; DNA; 1600 BP.
XX
XX AAV65786;
XX
XX 02-FEB-1999 (first entry)
DT
XX
XX Thermus aquaticus nuclease clone 3F (Cleavase BB) DNA.
DE
XX
XX Nucleic acid detection; multiple sequential invasive cleavage;
KW DNA polymerase; nuclease; Cleavase BB; ds.
XX
XX Thermus aquaticus strain YT-1.
OS Synthetic.
OS
XX
XX WO9842873-A1.
PN
XX
XX 01-OCT-1998.
PD
XX
XX 24-MAR-1998; 98WO-US05809.
PF
XX 24-MAR-1997; 97US-0823516.
PR
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

```

XX  Brow MAD, Hall JG, Kwiatkowski RW, Lyamichev VI;
PI  Mast AL, Vavra SH;
XX
XX  WPI; 1998-557036/47.
XX
XX  Detecting target nucleic acid by sequence-specific cleavage of
PT  complex with two specific oligonucleotides - used to detect
PT  cytomegalovirus DNA
XX
XX  Example 2; Page 278; 524pp; English.
XX
XX  This nucleotide sequence (clone 3F) codes for a thermostable
CC  nuclease (Cleavease BB) derived from the DNA polymerase (DNAP) of
CC  Thermus aquaticus (Taq). In comparison to the wild-type Taq DNAP
CC  sequence (see AAV65779), it contains an in-frame deletion of 903
CC  nucleotides. The invention relates to means for the detection and
CC  characterisation of nucleic acid sequences, and variations in
CC  nucleic acid sequences. It also relates to methods for forming a
CC  nucleic acid cleavage structure on a target sequence and cleaving
CC  this structure in a site-specific manner, preferably using a
CC  thermostable structure-specific nuclease such as a modified Taq
CC  DNAP that has reduced synthetic activity (see AAV65783-86). Cleavage
CC  of the cleavage structure by the nuclease indicates the presence of
CC  specific nucleic acid sequences or specific variants. The invention
CC  further relates to methods for the separation of nucleic acid
CC  molecules based on charge, methods for the detection of non-target
CC  cleavage products via the formation of a complete and activated
CC  protein binding region, and methods for the detection of nucleic
CC  acid from various viruses (e.g. human cytomegalovirus) in a sample.
CC  The method amplifies the detection molecule rather than the target
CC  itself, is less subject to contamination than exponential
CC  amplification processes, and allows many targets to be analysed in
CC  a single reaction.
XX
XX  Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
SQ

```

Alignment Scores:

Pred. No.:	1.89	Length:	1600
Score:	44.00	Matches:	9
Percent Similarity:	81.82%	Conservative:	0
Best Local Similarity:	81.82%	Mismatches:	2
Query Match:	93.62%	Indels:	0
DB:	19	Gaps:	0

```

US-09-823-649A-2 (1-11) x AAV65786 (1-1600)
QY  1 LeuSer***GluLeu***IleProTyGluGlu 11
DT  ||||| ||||| ||||| ||||| |||||
XX  Db 1133 CTCTCCAGAGAGCTAGCCATCCTTACGAGAG 1165

```

RESULT 6

AAV53855 standard; DNA: 1600 BP.

AAV53855;

21-DEC-1998 (first entry)

Nucleotide sequence of clone 4F of the Taq gene mutant.

Clone 4D; Taq mutant gene; thermostable; structure-specific nuclease;

mutant DNA polymerase; bacteria; fungi; protozoa; RNA virus;

hepatitis C virus; HCV; ds.

Thermus sp.

WO9823774-A1.

04-JUN-1998.

26-NOV-1997; 97WO-US21783.

```

PR  02-DEC-1996; 96US-0758314.
PR  29-NOV-1996; 96US-0757653.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
PI  Kaiser MW, Lyamichev VI, Lyamicheva N;
XX
XX  WPI; 1998-322748/28.
XX
XX  Thermostable structure-specific nuclease(s) derived from mutant DNA
PT  polymerase(s) - useful for detecting mutant allele(s) or strains of
PT  microorganisms
XX
XX  Example 2; Page 238; 472pp; English.
XX
XX  This is the nucleotide sequence of a clone of a mutant Taq gene,
CC  used in the method of the invention. In this process thermostable
CC  structure-specific nucleases are derived from mutant DNA polymerases,
CC  which can be used for detecting mutant alleles or strains of
CC  microorganisms. The structure-specific nucleases can be used in
CC  mixtures, compositions and kits to treat nucleic acid, e.g. for
CC  detection of wild type and mutant alleles of genes, for detection
CC  and/or identification of strains of microorganisms such as bacteria,
CC  fungi, protozoa, especially for detection of RNA viruses such as the
XX  hepatitis C virus (HCV).
XX
XX  Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
SQ

```

Alignment Scores:

Pred. No.:	1.89	Length:	1600
Score:	44.00	Matches:	9
Percent Similarity:	81.82%	Conservative:	0
Best Local Similarity:	81.82%	Mismatches:	2
Query Match:	93.62%	Indels:	0
DB:	19	Gaps:	0

```

US-09-823-649A-2 (1-11) x AAV53855 (1-1600)
QY  1 LeuSer***GluLeu***IleProTyGluGlu 11
DT  ||||| ||||| ||||| ||||| |||||
XX  Db 1133 CTCTCCAGAGAGCTAGCCATCCTTACGAGAG 1165

```

RESULT 7

AAV63407 standard; DNA: 1600 BP.

AAV63407;

26-JAN-1999 (first entry)

DNA sequence of a thermostable 5' nuclease derived from Taq polymerase.

Thermostable DNA polymerase; nucleic acid detection; Cleavease BB;

Thermostable 5' nuclease; ds.

Synthetic.

Thermus aquaticus.

US5837450-A.

17-NOV-1998.

06-JUN-1995; 95US-0471066.

06-JUN-1994; 94US-0254359.

07-DEC-1992; 92US-0986330.

04-JUN-1993; 93US-0073384.

06-JUN-1995; 95US-0471066.

(THIR-) THIRD WAVE TECHNOLOGIES INC.

Brow MAD, Dahlberg JE, Lyamichev VI.


```

DR WPI; 1999-023438/02.
XX
XX Detection of target nucleic acid molecules - uses modified
PT thermostable enzymes with specific cleavage activity to create
PT specific detection products from oligo-nucleotide(s) and target
PT hybridisation
XX
XX Claim 10; Columns 79-82; 91pp; English.
XX
CC The present sequence represents a thermostable 5' nuclease derived from
CC a thermostable polymerase modified to have reduced synthetic activity,
CC where the 5' nuclease is capable of cleaving a linear nucleic acid
CC duplex structure to create a single, single-stranded cleavage product.
CC The nuclease, designated cleavage BB, is used in a method for detecting
CC the presence of a nucleic acid molecule. The method is used for the
CC specific detection of nucleic acid sequences, via a cleavage-based
CC procedure, but without the need for amplification of target sequences.
CC Thermostable polymerases, altered to have nuclease, but not polymerase
CC activity are preferably used due to their specificity. The cleavage
CC product specifically formed is detected, preferably by the use of
CC radioactively labelled oligonucleotides. These can be used in
CC e.g. forensic testing or paternity determination.
XX
SQ Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.89 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 20 Gaps: 0
XX
US-09-823-649A-2 (1-11) x AAV63407 (1-1600)
QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
DB 1133 CTCCTCCAGAGCTAGCCATCCCTTAGCAGAGAG 1165
XX
RESULT 8
ID AAQ24010 standard; DNA; 1635 BP.
XX
AC AAQ24010;
XX
XX 22-OCT-1992 (first entry)
XX
DE Mutant thermostable DNA polymerase pSPSA286.
XX
XX 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCK; ss.
XX
OS Thermus species sp17.
XX
XX
XX Key Location/Qualifiers
XX old_sequence 1..2
XX /*tag= a
XX /*note= "nucleotides 4-861 deleted from the native
XX sequence."
XX
XX
XX MO9206200-A.
XX
XX 16-APR-1992.
XX
XX 30-SEP-1991; 91MO-US07035.
XX
XX 28-SEP-1990; 90US-0590213.
XX 28-SEP-1990; 90US-0590466.
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU ) CETUS CORP.
XX
XX Abramson RD, Gelfand DH.
XX

```

```

DR WPI; 1992-150885/18.
DR P-PSDB; AAR23157.
XX
XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
XX Claim 11; Page 59; 185pp; English.
XX
CC The sequence is that of Thermus species sp17 polymerase DNA which
CC has been mutated. The mutation designated pSPSA286 causes the
CC polymerase enzyme produced to exhibit a different aml. of 5'-3'
CC exonuclease activity than the native enzyme. Thermostable DNA
CC polymerases are useful in many recombinant DNA techniques, esp.
CC nucleic acid amplification by PCR, self-sustained sequence
CC replication (SSR) and high temp. DNA sequencing. The absence of
CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
CC discrimination in a combined polymerase ligase chain reaction (PLCR)
CC assay. An enhanced aml. of 5'-3' exonuclease activity may be desirable
CC in enzymes used in homogeneous assays for the amplification and
CC detection of a target nucleic acid sequence. Mutation of the DNA
CC encoding particular regions of the enzymes can be used to prepare
CC a range of recombinant proteins having 5'-3' exonuclease activity
CC to a complete lack of activity.
XX
XX See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
SQ Sequence 1635 BP; 267 A; 572 C; 559 G; 237 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.93 Length: 1635
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 13 Gaps: 0
XX
US-09-823-649A-2 (1-11) x AAQ24010 (1-1635)
QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
DB 1168 CTCCTCCAGAGCTCTCCATCCCTTAGCAGAG 1200
XX
RESULT 9
ID AAQ23998
XX
AC AAQ23998 standard; DNA; 1635 BP.
XX
XX 22-OCT-1992 (first entry)
XX
DE Mutant thermostable DNA polymerase p15CB.
XX
XX 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCK; ss.
XX
XX
XX Thermus aquaticus.
XX
XX
XX Key Location/Qualifiers
XX old_sequence 1..2
XX /*tag= a
XX /*note= "nucleotides 4-867 deleted from native
XX sequence"
XX
XX
XX MO9206200-A.
XX
XX 16-APR-1992.
XX
XX 30-SEP-1991; 91MO-US07035.
XX
XX 28-SEP-1990; 90US-0590213.
XX 28-SEP-1990; 90US-0590466.
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU ) CETUS CORP.
XX
XX

```

XX
PI Abramson RD, Gelfand DH;
XX
DR WPI: 1992-150865/18.
DR P-PSDB; AAR23145.
XX
PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
PS Claim 11; Page 59; 185pp; English.
XX
CC The sequence is that of *Thermus aquaticus* polymerase DNA which has
CC been mutated. The mutation, resulting in mutant pL5G8, causes
CC the polymerase enzyme produced to exhibit a different amt. of 5'-3'
CC exonuclease activity than the native enzyme. Thermostable DNA
CC polymerases are useful in many recombinant DNA techniques, esp.
CC nucleic acid amplification by PCR, self-sustained sequence
CC replication (SSR) and high temp. DNA sequencing. The absence of 5'-3'
CC nuclease activity may facilitate higher sensitivity allelic
CC discrimination in a combined polymerase ligase chain reaction (PLCR)
CC assay. An enhanced amt. of 5'-3' exonuclease activity may be
CC desirable in enzymes used in homogeneous assays for the amplification
CC and detection of a target nucleic acid sequence. Mutation of the DNA
CC encoding particular regions of the enzymes can be used to prepare a
CC range of recombinant proteins having 5'-3' exonuclease activity to a
CC complete lack of activity.
CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
SQ Sequence 1635 BP; 274 A; 551 C; 564 G; 246 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.93 * Length: 1635
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 13 Gaps: 0
US-09-823-649A-2 (1-11) x AAQ23998 (1-1635)
QY 1 LeuSer***Glueu***IleProTyrglu 11
DB 1168 CTTCTCCAGAGCTACGCCATCCCTTACGAGSAG 1200
RESULT 10
AAQ24322
ID AAQ24322 standard; DNA; 1635 BP.
XX
AC AAQ24322;
XX
DT 22-OCT-1992 (first entry)
XX
DE Mutant thermostable DNA polymerase pZ05A292.
XX
KM 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
XX
OS *Thermus* species Z05.
XX
FH Key Location/Qualifiers
FT old_sequence 1..2
FT /*tag= a
FT /note= "nucleotides 4-873 deleted from the native
FT sequence."
XX
PN WO9206200-A.
XX
PD 16-APR-1992.
XX
PF 30-SEP-1991; 91WO-US07035.
XX
PR 28-SEP-1990; 90US-0590213.
PR 28-SEP-1990; 90US-0590466.

PR 28-SEP-1990; 90US-0590490.
XX
PA (CETU) CETUS CORP.
XX
PI Abramson RD, Gelfand DH;
XX
DR WPI: 1992-150865/18.
DR P-PSDB; AAR23163.
XX
PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
PS Claim 11; Page 59; 185pp; English.
XX
CC The sequence is that of *Thermus* species Z05 polymerase DNA which
CC has been mutated. The mutation designated pZ05A292 causes the
CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
CC exonuclease activity than the native enzyme. Thermostable DNA
CC polymerases are useful in many recombinant DNA techniques, esp.
CC nucleic acid amplification by PCR, self-sustained sequence
CC replication (SSR) and high temp. DNA sequencing. The absence of
CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
CC discrimination in a combined polymerase ligase chain reaction (PLCR)
CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
CC in enzymes used in homogeneous assays for the amplification and
CC detection of a target nucleic acid sequence. Mutation of the DNA
CC encoding particular regions of the enzymes can be used to prepare
CC a range of recombinant proteins having 5'-3' exonuclease activity
CC to a complete lack of activity.
CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
SQ Sequence 1635 BP; 291 A; 577 C; 529 G; 238 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.93 Length: 1635
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 13 Gaps: 0
US-09-823-649A-2 (1-11) x AAQ24322 (1-1635)
QY 1 LeuSer***Glueu***IleProTyrglu 11
DB 1168 CTTCTCCAGAGCTACGCCATCCCTTACGAGSAG 1200
RESULT 11
AAQ24328
ID AAQ24328 standard; DNA; 1635 BP.
XX
AC AAQ24328;
XX
DT 22-OCT-1992 (first entry)
XX
DE Mutant thermostable DNA polymerase pYTHA292.
XX
KM 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
XX
OS *Thermus* thermophilus.
XX
FH Key Location/Qualifiers
FT old_sequence 1..2
FT /*tag= a
FT /note= "nucleotides 4-873 deleted from the native
FT sequence."
XX
PN WO9206200-A.
XX
PD 16-APR-1992.
XX
PF 30-SEP-1991; 91WO-US07035.

XX 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CERTU) CERTUS CORP.
 XX
 PI Abramson RD, Gelfand DH;
 XX
 DR WPI: 1992-150885/18.
 DR P-PSDB; AAR23168.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 PS
 PS Claim 11; Page 59; 185pp: English.
 XX
 CC The sequence is that of Thermus thermophilus polymerase DNA which
 CC has been mutated. The mutation designated PTNHA292 causes the
 CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. Thermostable DNA
 CC polymerases are useful in many recombinant DNA techniques, esp.
 CC nucleic acid amplification by PCR, self-sustained sequence
 CC replication (SSR) and high temp. DNA sequencing. The absence of
 CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
 XX
 SQ Sequence 1635 BP; 283 A; 577 C; 238 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.93 Length: 1635
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 13 Gaps: 0
 US-09-823-649A-2 (1-11) x AAQ24328 (1-1635)
 QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
 DB 1168 CTCCTCCAGAGCTTGCACCTCCACGAGAG 1200
 RESULT 12
 AA229460
 ID AA229460 standard; DNA; 1682 BP.
 XX
 AC AA229460;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Thermostable DNA dependent mutant Taq polymerase-1 encoding DNA.
 XX
 KW Thermostable DNA dependent polymerase; nuclease activity; Taq polymerase;
 KW mutant; Polymerase Chain Reaction; PCR; N-terminal domain; proteolytic;
 KW thermal stability; hydrophilic solution; sequencing; labeling; diagnosis;
 KW genetic disorder; primer-driven mutagenesis; identification of pathogen;
 KW mutational analysis; forensic identification; ds.
 XX
 OS Thermus aquaticus.
 XX
 FH Key Location/Qualifiers
 FT CDS 8..1669
 FT /*tag= a

PT
 FT
 FT
 FT
 XX
 XX
 PN W0964438-A1.
 XX
 PD 16-DEC-1999.
 XX
 PE 11-JUN-1999; 99WO-US13305.
 XX
 PR 11-JUN-1998; 98US-0096399.
 XX
 PA (CLON-) CLONTECH LAB INC.
 XX
 PI Wurst H, Qui Z;
 XX
 DR WPI: 2000-105869/09.
 DR P-PSDB; AAY44352.
 XX
 PT Novel enzyme, particularly useful in polymerase chain reaction
 PT
 PS Claim 22; Fig 1; 32pp: English.
 XX
 CC The present sequence is the DNA encoding a novel, thermostable DNA
 CC dependent, mutant Taq polymerase-1 enzyme. The N-terminal domain of the
 CC polymerase includes a sequence of 9 amino acids, that has 40-50%
 CC sequence identity with residues 280-288 of native Thermus aquaticus
 CC polymerase, Taq. The residues 10-553 of encoded by this sequence is
 CC identical to residues 289-832 of Taq polymerase. This sequence has no
 CC significant nuclease activity, but good proteolytic and thermal
 CC stabilities and improved solubility in hydrophilic solutions. This
 CC sequence is used for synthesis of polynucleotides, particularly in
 CC polymerase chain reaction (PCR) based processes, like sequencing,
 CC labeling, primer-driven mutagenesis, diagnosis of genetic disorders,
 CC identification of pathogens, mutational analysis and forensic
 CC identification.
 XX
 SQ Sequence 1682 BP; 286 A; 567 C; 575 G; 254 T; 0 other;
 Alignment Scores:
 Pred. No.: 2 Length: 1682
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 21 Gaps: 0
 US-09-823-649A-2 (1-11) x AA229460 (1-1682)
 QY 1 LeuSer***GluLeu***IleProTyrGluGlu 11
 DB 1202 CTCCTCCAGAGCTAGCCATCCCTTAGAGGAG 1234
 RESULT 13
 AA147959
 ID AA147959 standard; DNA; 1686 BP.
 XX
 AC AA147959;
 XX
 DT 11-MAY-1997 (first entry)
 XX
 DE Mutant Taq polymerase FY2.
 XX
 KW Tag; polymerase; FY2; FY3; FY4; thermostable; Thermus; enzyme;
 KW DNA sequencing; PCR; ss.
 XX
 OS Thermus aquaticus.
 XX
 PN EP745676-A1.
 XX
 PD 04-DEC-1996.
 XX
 PF 30-MAY-1996; 96EP-0303880.

/product= "Thermostable mutant Taq polymerase-1"
 /note= "Identical at C-terminus to Taq polymerase
 from Thermus aquaticus"

XX 31-MAY-1995; 95US-0455686.
 XX (AMSH) AMERSHAM LIFE SCI INC.
 XX Davis M, Fuller C, Moffett B, Fuller CW, Moffett RB;
 XX WPI: 1997-013699/02.
 XX P-PSDB: AAM09315.
 XX
 PT New enzymatically active *Thermus* DNA polymerase(s) - having a
 PT tyrosine at a position equivalent to 667 of *Taq* DNA polymerase and
 PT lacking 5' to 3' exonuclease activity
 XX
 PS Claim 5; Page 10-13; 36pp; English.
 XX
 CC FY2, FY3 and FY4 (see AAT47959 to AAT47961) are new thermostable DNA
 CC polymerases providing improved results in DNA sequencing reactions.
 CC FY2 is the *T. aquaticus* DNA polymerase lacking pref. the 271-272
 CC N-terminal amino acids and having the amino acid at position 667 of
 CC the wild type (Phe) replaced by Tyr. FY3 contains 562 amino acids
 CC with a Met at position 1 and Ala at position 2, corresponding to the
 CC Met and Ala of positions 271 and 272, respectively, of the wild type
 CC enzyme. FY4 corresponds to the *T. thermophilus* DNA polymerase
 CC lacking the N-terminal 274 amino acids having a Phe to Tyr
 CC replacement at position 396 (position 669 of the wild type enzyme).
 XX
 SQ Sequence 1686 BP; 283 A; 567 C; 578 G; 258 T; 0 other:
 XX
 Alignment Scores:
 Pred. No.: 2 Length: 1686
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 18 Gaps: 0
 US-09-823-649a-2 (1-11) x AAT47959 (1-1686)
 QY 1 Leuser***Glueu***IleProTyrGlnGlu 11
 DB 1219 CTCCTCCAGAGCTAGCCATCCCTACGAGAG 1251
 RESULT 14
 AAT47961
 ID AAT47961 standard; DNA; 1686 BP.
 XX
 AC AAT47961;
 XX
 DT 11-MAY-1997 (first entry)
 XX
 DE Mutant *Taq* polymerase FY4.
 XX
 KW Tag: polymerase; FY2; FY3; FY4; thermostable; *Thermus*; enzyme;
 KW DNA sequencing; PCR; ss.
 XX
 OS *Thermus thermophilus*.
 XX
 PN EP745676-A1.
 XX
 PD 04-DEC-1996.
 XX
 PF 30-MAY-1996; 96EP-0303880.
 XX
 PR 31-MAY-1995; 95US-0455686.
 XX
 PA (AMSH) AMERSHAM LIFE SCI INC.
 XX
 PI Davis M, Fuller C, Moffett B, Fuller CW, Moffett RB;
 XX
 DR WPI: 1997-013699/02.
 DR P-PSDB: AAM09317.
 XX

PT New enzymatically active *Thermus* DNA polymerase(s) - having a
 PT tyrosine at a position equivalent to 667 of *Taq* DNA polymerase and
 PT lacking 5' to 3' exonuclease activity
 XX
 PS Claim 5; Page 20-23; 36pp; English.
 XX
 CC FY2, FY3 and FY4 (see AAT47959 to AAT47961) are new thermostable DNA
 CC polymerases providing improved results in DNA sequencing reactions.
 CC FY2 is the *T. aquaticus* DNA polymerase lacking pref. the 271-272
 CC N-terminal amino acids and having the amino acid at position 667 of
 CC the wild type (Phe) replaced by Tyr. FY3 contains 562 amino acids
 CC with a Met at position 1 and Ala at position 2, corresponding to the
 CC Met and Ala of positions 271 and 272, respectively, of the wild type
 CC enzyme. FY4 corresponds to the *T. thermophilus* DNA polymerase
 CC lacking the N-terminal 274 amino acids having a Phe to Tyr
 CC replacement at position 396 (position 669 of the wild type enzyme).
 XX
 SQ Sequence 1686 BP; 295 A; 593 C; 551 G; 247 T; 0 other:
 XX
 Alignment Scores:
 Pred. No.: 2 Length: 1686
 Score: 44.00 Matches: 9
 Percent Similarity: 81.82% Conservative: 0
 Best Local Similarity: 81.82% Mismatches: 2
 Query Match: 93.62% Indels: 0
 DB: 18 Gaps: 0
 US-09-823-649a-2 (1-11) x AAT47961 (1-1686)
 QY 1 Leuser***Glueu***IleProTyrGlnGlu 11
 DB 1219 CTCCTCCAGAGCTAGCCATCCCTACGAGAA 1251
 RESULT 15
 AAX27130
 ID AAX27130 standard; DNA; 1686 BP.
 XX
 AC AAX27130;
 XX
 DT 27-MAY-1999 (first entry)
 XX
 DE FY2 polymerase coding sequence.
 XX
 KW FY2 polymerase; FY3 polymerase; FY4 polymerase; *Taq* polymerase variant;
 KW DNA sequencing; ss.
 XX
 OS *Thermus aquaticus*.
 XX
 PN US5885813-A.
 XX
 PD 23-MAR-1999.
 XX
 PF 14-MAY-1996; 96US-0648657.
 XX
 PR 14-MAY-1996; 96US-0648657.
 PR 31-MAY-1995; 95US-0455686.
 XX
 PA (AMSH) AMERSHAM LIFE SCI INC.
 XX
 PI Cunliff JJ, Davis M, Fuller CW, Moffett RB;
 XX
 DR WPI: 1999-228538/19.
 DR P-PSDB: AAY00886.
 XX
 PT New thermostable DNA polymerases - having 540 to 582 amino acids
 PT and lacking 5' to 3' exonuclease activity
 XX
 PS Claim 7; Column 11-16; 27pp; English.
 XX
 CC This sequence encodes a thermostable DNA polymerase of the invention. The
 CC thermostable polymerases of the invention are variants of *Taq* DNA
 CC polymerase having a Tyr residue at a position corresponding to *Taq* DNA
 CC polymerase residue 667 in its dNMP binding site, and lacking 5' to 3'

CC exonuclease activity as a result of an N-terminal deletion. The
CC polymerases can be used for the sequencing of DNA, either by manual or
CC automated means.

XX
SQ Sequence 1686 BP; 283 A; 567 C; 578 G; 258 T; 0 other;

Alignment Scores:

Pred. No.:	2	Length:	1686
Score:	44.00	Matches:	9
Percent Similarity:	81.82%	Conservative:	0
Best Local Similarity:	81.82%	Mismatches:	2
Query Match:	93.62%	Indels:	0
DB:	20	Gaps:	0

US-09-823-649A-2 (1-11) x AAX27130 (1-1686)

QY 1 LeuSer**GIuLeu**IIeProTYrGIuGlu 11

||||| ||||| ||||| ||||| |||||

Db 1219 CTCTCCAGAGACTAGCCATCCCTTACGAGAG 1251

Search completed: January 15, 2003, 12:55:25
Job time : 135.143 secs

PT using a mutant thermoactive DNA polymerase -
XX Claim 4; Page 3; 23pp; English.
XX
XX The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction. This motif is derived from DNA polymerases from
CC Thermus species aquaticus, thermophilus, 205 and caldophilus.
XX
SQ Sequence 11 AA;
Query Match 93.6%; Score 44; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 LSXELXIPYEE 11
| | | | |
Db 1 LSOELAIPIYEE 11
- - -
RESULT 2
AAB47794
ID AAB47794 standard; peptide; 11 AA.
XX
XX AAB47794;
XX
XX 25-MAR-2002 (first entry)
XX
XX Native DNA polymerase motif #4.
XX
XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KW transverse transcription reaction; fluorescein; cyanine; thermoactive;
KW dye; amplification.
XX
XX Thermus sp.
OS
XX
XX Key Location/Qualifiers
FH Misc-difference 3
FT /Label= Gln, Gly
XX
XX EPIJ52062-A2.
XX
XX 07-NOV-2001.
XX
XX 12-APR-2001; 2001EP-0109341.
XX
XX 18-APR-2000; 2000US-198336P.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
XX Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
XX
XX WPI; 2002-076891/11.
XX
XX Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
XX using a mutant thermoactive DNA polymerase -
XX
XX Claim 4; Page 4; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction. This motif is derived from DNA polymerases
CC from Thermus species flavus, sp017 and filliformis.
XX
SQ Sequence 11 AA;
Query Match 93.6%; Score 44; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LSXELXIPYEE 11
| | | | |
Db 1 LSOELXIPYEE 11
- - -
RESULT 3
AAM48259
ID AAM48259 standard; Peptide; 11 AA.
XX
XX AAM48259;
XX
XX 25-MAR-2002 (first entry)
XX
XX Native DNA polymerase motif #8.
XX
XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KW transverse transcription reaction; fluorescein; cyanine; thermoactive;
KW dye; amplification.
XX
XX Thermus aquaticus.
OS
XX
XX EPIJ52062-A2.
XX
XX 07-NOV-2001.
XX
XX 12-APR-2001; 2001EP-0109341.
XX
XX 18-APR-2000; 2000US-198336P.
XX
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
XX Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
XX
XX WPI; 2002-076891/11.
XX
XX Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
XX using a mutant thermoactive DNA polymerase -
XX
XX Disclosure; Page 7; 23pp; English.
XX
XX The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is

CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

CC Sequence 11 AA;

Query Match 93.6%; Score 44; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSKELXIPYEE 11
1111111111
DB 1 LSGELAIPIYEE 11

RSULT 4

AA48260

ID AAM48260 standard; Peptide: 11 AA.

AC AAM48260;

DF 25-MAR-2002 (first entry)

DE Native DNA polymerase motif #9.

XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;

KW transverse transcription reaction; fluorescein; cyanine; thermoactive;

KW dye; amplification.

XX Thermus flavus.

OS Ep1152062-A2.

PN 07-NOV-2001.

XX 12-APR-2001; 2001EP-0109341.

XX 18-APR-2000; 2000US-198336P.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;

PI Schoenbrunner NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermoactive DNA polymerase

PS Disclosure; Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

SO Sequence 11 AA;

Query Match 93.6%; Score 44; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSKELXIPYEE 11
1111111111
DB 1 LSGELAIPIYEE 11

RESULT 5

AA48261

ID AAM48261 standard; Peptide: 11 AA.

AC AAM48261;

DF 25-MAR-2002 (first entry)

DE Native DNA polymerase motif #10.

XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;

KW transverse transcription reaction; fluorescein; cyanine; thermoactive;

KW dye; amplification.

OS Thermus thermophilus.

PN Ep1152062-A2.

XX 07-NOV-2001.

XX 12-APR-2001; 2001EP-0109341.

XX 18-APR-2000; 2000US-198336P.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;

PI Schoenbrunner NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermoactive DNA polymerase

PS Disclosure; Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

SO Sequence 11 AA;

Query Match 93.6%; Score 44; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSKELXIPYEE 11
1111111111
DB 1 LSGELAIPIYEE 11

RESULT 6

AAAM8262 standard; Peptide; 11 AA.

AC AAAM8262;

DT 25-MAR-2002 (first entry)

DE Native DNA polymerase motif #11.

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
 KW transverse transcription reaction; fluorescein; cyanine; thermoactive;
 KW dye; amplification.

OS Thermus species Z05.

PN EPI152062-A2.

PD 07-NOV-2001.

PF 12-APR-2001; 2001EP-0109341.

PR 18-APR-2000; 2000US-198336P.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
 PI Schoenbrunner NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
 PT transcriptase polymerase chain reaction amplification of a mixture
 PT using a mutant thermoactive DNA polymerase -

PS Disclosure; Page 7; 23pp; English.

CC The sequences given in AAM47791-97 and AAM48259-AAAM48270 represent the
 CC native forms of motifs derived from DNA polymerases used in the method
 CC of the invention. The method for reverse transcribing an RNA, comprises
 CC treating a transverse transcription reaction mixture which comprises the
 CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
 CC polymerase. These motifs represent a conserved region which affects the
 CC DNA polymerase's ability to incorporate deoxynucleotides labelled with
 CC fluorescein and cyanine family dyes. The method of the invention is
 CC useful in reverse transcription/amplification reactions. The method
 CC provides improved reverse transcription efficiency at lower enzyme
 CC concentrations relative to previous high temperature reverse
 CC transcription methods, and provide a wider range of usable salt
 CC concentrations. The use of the mutant DNA polymerase provides faster
 CC reverse transcription extension rates, and consequently less time is
 CC needed for the reaction.

SQ Sequence 11 AA;

Query Match 93.6%; Score 44; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11

DB 1 LSOELAIPIYEE 11

RESULT 7

AAAM8263 standard; Peptide; 11 AA.

AC AAAM8263;

DT 25-MAR-2002 (first entry)

XX

DE Native DNA polymerase motif #12.

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
 KW transverse transcription reaction; fluorescein; cyanine; thermoactive;
 KW dye; amplification.

OS Thermus species sps17.

PN EPI152062-A2.

PD 07-NOV-2001.

PF 12-APR-2001; 2001EP-0109341.

PR 18-APR-2000; 2000US-198336P.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
 PI Schoenbrunner NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
 PT transcriptase polymerase chain reaction amplification of a mixture
 PT using a mutant thermoactive DNA polymerase -

PS Disclosure; Page 7; 23pp; English.

CC The sequences given in AAM47791-97 and AAM48259-AAAM48270 represent the
 CC native forms of motifs derived from DNA polymerases used in the method
 CC of the invention. The method for reverse transcribing an RNA, comprises
 CC treating a transverse transcription reaction mixture which comprises the
 CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
 CC polymerase. These motifs represent a conserved region which affects the
 CC DNA polymerase's ability to incorporate deoxynucleotides labelled with
 CC fluorescein and cyanine family dyes. The method of the invention is
 CC useful in reverse transcription/amplification reactions. The method
 CC provides improved reverse transcription efficiency at lower enzyme
 CC concentrations relative to previous high temperature reverse
 CC transcription methods, and provide a wider range of usable salt
 CC concentrations. The use of the mutant DNA polymerase provides faster
 CC reverse transcription extension rates, and consequently less time is
 CC needed for the reaction.

SQ Sequence 11 AA;

Query Match 93.6%; Score 44; DB 23; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11

DB 1 LSOELAIPIYEE 11

RESULT 8

AAAM8264 standard; Peptide; 11 AA.

AC AAAM8264;

DT 25-MAR-2002 (first entry)

DE Native DNA polymerase motif #13.

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
 KW transverse transcription reaction; fluorescein; cyanine; thermoactive;
 KW dye; amplification.

OS Thermus caldophilus.

PN EPI152062-A2.


```

XX PD 07-NOV-2001.
XX PF 12-APR-2001; 2001EP-0109341.
XX PR 18-APR-2000; 2000US-198336P.
XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
XX PI Schoenbrunner NJ, Wang AM;
XX DR WPI; 2002-076891/11.
XX PT Reverse transcribing an RNA, comprises performing a reverse
XX PT transcriptase polymerase chain reaction amplification of a mixture
XX PT using a mutant thermocactive DNA polymerase
XX PS Disclosure; Page 7; 23pp; English.
XX CC The sequences given in AAM47791-97 and AAM48259-AAM48270 represent the
XX CC native forms of motifs derived from DNA polymerases used in the method
XX CC of the invention. The method for reverse transcribing an RNA, comprises
XX CC treating a transverse transcription reaction mixture which comprises the
XX CC RNA, a primer, a divalent cation, and a mutant thermocactive DNA
XX CC polymerase. These motifs represent a conserved region which affects the
XX CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
XX CC fluorescein and cyanine family dyes. The method of the invention is
XX CC useful in reverse transcription/amplification reactions. The method
XX CC provides improved reverse transcription efficiency at lower enzyme
XX CC concentrations relative to previous high temperature reverse
XX CC transcription methods, and provide a wider range of usable salt
XX CC concentrations. The use of the mutant DNA polymerase provides faster
XX CC reverse transcription extension rates, and consequently less time is
XX CC needed for the reaction.
XX SQ Sequence 11 AA:
XX
XX Query Match 93.6%; Score 44; DB 23; Length 11;
XX Best Local Similarity 81.8%; Pred. No. 0.011;
XX Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 LSXELXIPYE 11
XX 11 11 11111
XX DB 1 LSQELXIPYE 11
XX
XX RESULT 9
XX AAM48265
XX ID AAM48265 standard; Peptide; 11 AA.
XX AC AAM48265;
XX XX
XX DT 25-MAR-2002 (first entry)
XX XX
XX DE Native DNA polymerase motif #14.
XX XX
XX KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
XX KW transverse transcription reaction; fluorescein; cyanine; thermocactive;
XX KW dye; amplification.
XX XX
XX OS Thermus filiformis.
XX PN EP1152062-A2.
XX XX
XX PD 07-NOV-2001.
XX PF 12-APR-2001; 2001EP-0109341.
XX PR 18-APR-2000; 2000US-198336P.
XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX

```

```

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
XX DR WPI; 2002-076891/11.
XX PT Reverse transcribing an RNA, comprises performing a reverse
XX PT transcriptase polymerase chain reaction amplification of a mixture
XX PT using a mutant thermocactive DNA polymerase
XX PS Disclosure; Page 7; 23pp; English.
XX CC The sequences given in AAM47791-97 and AAM48259-AAM48270 represent the
XX CC native forms of motifs derived from DNA polymerases used in the method
XX CC of the invention. The method for reverse transcribing an RNA, comprises
XX CC treating a transverse transcription reaction mixture which comprises the
XX CC RNA, a primer, a divalent cation, and a mutant thermocactive DNA
XX CC polymerase. These motifs represent a conserved region which affects the
XX CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
XX CC fluorescein and cyanine family dyes. The method of the invention is
XX CC useful in reverse transcription/amplification reactions. The method
XX CC provides improved reverse transcription efficiency at lower enzyme
XX CC concentrations relative to previous high temperature reverse
XX CC transcription methods, and provide a wider range of usable salt
XX CC concentrations. The use of the mutant DNA polymerase provides faster
XX CC reverse transcription extension rates, and consequently less time is
XX CC needed for the reaction.
XX SQ Sequence 11 AA:
XX
XX Query Match 93.6%; Score 44; DB 23; Length 11;
XX Best Local Similarity 81.8%; Pred. No. 0.011;
XX Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 LSXELXIPYE 11
XX 11 11 11111
XX DB 1 LSQELXIPYE 11
XX
XX RESULT 10
XX AAR96267
XX ID AAR96267 standard; Protein; 528 AA.
XX AC AAR96267;
XX XX
XX DT 18-NOV-1996 (first entry)
XX XX
XX DE Mutant Thermus aquaticus DNA polymerase (clone 4F).
XX XX
XX KW p53; mutant; mutation; cleavage; nuclease; cleavage; Thermus;
XX KW Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
XX KW Staphylococcus; Bacillus; detection.
XX XX
XX OS Thermus aquaticus.
XX PN W09615267-A1.
XX XX
XX PD 23-MAY-1996.
XX XX
XX PF 09-NOV-1995; 95WO-US14671.
XX XX
XX PR 30-AUG-1995; 95US-0520946.
XX PR 09-NOV-1994; 94US-0337164.
XX PR 09-MAR-1995; 95US-0402601.
XX PR 07-JUN-1995; 95US-0484956.
XX XX
XX PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX XX
XX PI Brow MAD, Dahlberg JE, Fors L, Heisler LM, Lyamichev VI;
XX PI Oldenburg MC, Olive DM;
XX DR WPI; 1996-259862/26.
XX DR N-PSDB; AAT27686.
XX

```

PT esp. in human p53 gene, to identify strains of microorganisms and
 PT viruses
 XX
 PS Example 2A, Page 287-288; 433pp; English.
 XX
 CC Cleavage of nucleic acids using an enzyme, especially a nuclease
 CC selected from the group consisting of cleavage (RII) BN enzyme,
 CC Thermus aquaticus DNA polymerase, Thermus thermophilus DNA
 CC polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
 CC Rad1/Rad10 complex. The nucleic acid substrate is preferably an
 CC oligonucleotide containing a human p53 gene sequence or
 CC alternatively, microbial gene sequences. Cleavage products are
 CC compared to the cleavage products of reference gene sequences. The
 CC method is used for detecting mutation in the human p53 gene; for
 CC identifying strains of microorganisms, especially bacteria selected
 CC from the group of members of the genera Campylobacter,
 CC Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
 CC The method may also be used for the identification of viruses,
 CC especially hepatitis C virus and simian immunodeficiency virus.
 CC Thermus aquaticus (Taq) DNA polymerase was amplified using two
 CC primers (AAT27679, AAT27680). The Taq polymerase DNA was inserted into
 CC the BamHI restriction site of the expression vector pET-3c and mutant
 CC genes were created from that construct. This mutant was created
 CC after the vector was digested with BstXI and BamHI. The DNA was
 CC then treated with the Klenow fragment of DNAseI to trim both 3'
 CC overhangs to blunt ends which were then ligated together, resulting
 CC in an in frame deletion of 903 nucleotides. This mutant Taq
 CC polymerase is also referred to as the cleavage BB enzyme.
 XX
 SQ Sequence 528 AA;
 XX
 Query Match 93.6%; Score 44; DB 17; Length 528;
 Best Local Similarity 81.8%; Pred. No. 0.86;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSXELXIPYEE 11
 DB 374 LSQELAIPIYEE 384
 XX
 RESULT 11
 AAR23145
 ID AAR23145 standard; Protein; 544 AA.
 XX
 AC AAR23145;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase enzyme MET-SER 290 Taq.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR.
 XX
 OS Thermus aquaticus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..2 /note="deletion of residues 2-289 of native
 FT sequence."
 XX
 PN WO9206200-A.
 XX
 PD 16-APR-1992.
 XX
 PE 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelfand DH;

... WPI; 1992-150885/18.
 DR N-PSDB; AAQ243998.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC The sequence is that of a mutant of Thermus aquaticus polymerase
 CC mutant MET-SER 290 Taq (the Stoffel fragment) having a different
 CC amt. of 5'-3' exonuclease activity than the native enzyme.
 CC Thermostable DNA polymerases are useful in many recombinant DNA
 CC techniques, esp. nucleic acid amplification by PCR, self-sustained
 CC sequence replication (SSR) and high temp. DNA sequencing. The absence
 CC of 5'-3' nuclease activity may facilitate higher sensitivity allele
 CC discrimination in a combined polymerase ligase chain reaction (PCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAR23140-79 and AAR23722.
 XX
 SQ Sequence 544 AA;
 XX
 Query Match 93.6%; Score 44; DB 13; Length 544;
 Best Local Similarity 81.8%; Pred. No. 0.89;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSXELXIPYEE 11
 DB 390 LSQELAIPIYEE 400
 XX
 RESULT 12
 AAR23163
 ID AAR23163 standard; Protein; 544 AA.
 XX
 AC AAR23163;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase enzyme MET-ALA 292 T205.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR.
 XX
 OS Thermus species Z05.
 XX
 PN WO9206200-A.
 XX
 PD 16-APR-1992.
 XX
 PE 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelfand DH;
 XX
 DR WPI; 1992-150885/18.
 DR N-PSDB; AAQ243322.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.

CC	are useful in many recombinant DNA techniques, esp. nucleic acid
CC	amplification by PCR, self-sustained sequence replication (SSR)
CC	and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
CC	may facilitate higher sensitivity allelic discrimination in a combined
CC	polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
CC	5'-3' exonuclease activity may be desirable in enzymes used in
CC	homogeneous assays for the amplification and detection of a target
CC	nucleic acid sequence. Mutation of the DNA encoding particular
CC	regions of the enzymes can be used to prepare a range of recombinant
CC	proteins having 5'-3' exonuclease activity to a complete lack of
CC	activity.
CC	See also AAR23140-79 and AAR23722.
XX	
SQ	Sequence 544 AA;
Db	
QY	1 LSNELXIPYEE 11
Db	390 LSQELAIPYEE 400
RESULT 14	
AAR23157	ID AAR23157 standard; Protein; 545 AA.
XX	
AC	AAR23157;
XX	
DT	22-OCT-1992 (first entry)
XX	
PB	Mutant thermostable DNA polymerase enzyme MET-ALA 288 Tpsp17.
XX	
KW	5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
XX	
OS	Thermus species SPS17.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 1..2
FT	/note= "residues 2-287 deleted from the native
FT	sequence"
XX	
PN	WO9206200-A.
XX	
ED	16-APR-1992.
XX	
PF	30-SEP-1991; 91WO-US07035.
XX	
PR	28-SEP-1990; 90US-0590213.
PR	28-SRP-1990; 90US-0590466.
PR	28-SRP-1990; 90US-0590490.
PA	(CETU) CETUS CORP.
XX	
PI	Adramson RD, Gelfand DH;
XX	
DR	WP1: 1992-150885/18.
XX	
N-P	PSDB; AAQ24010.
XX	
XX	
PT	Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT	activity - having conserved regions mutated or deleted, for use
PT	in e.g. PCR, sequencing and detection assays
PS	
Claim 11:	Page 59; 185pp; English.
CC	The sequence is that of a mutant of Thermus species sps17 polymerase
CC	mutant MET-ALA 288 Tpsp17, having a different amt. of 5'-3' exonuclease
CC	activity than the native enzyme. Thermostable DNA polymerases
CC	are useful in many recombinant DNA techniques, esp. nucleic acid
CC	amplification by PCR, self-sustained sequence replication (SSR)
CC	and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
CC	may facilitate higher sensitivity allelic discrimination in a combined

CC polymerase ligase chain reaction (PCR) assay. An enhanced amt. of
CC 5'-3' exonuclease activity may be desirable in enzymes used in
CC homogeneous assays for the amplification and detection of a target
CC nucleic acid sequence. Mutation of the DNA encoding particular
CC regions of the enzymes can be used to prepare a range of recombinant
CC proteins having 5'-3' exonuclease activity to a complete lack of
CC activity.
CC See also AAR23140-79 and AAR23722.

```

QY      1 LSXELXIPYEE 11
          || || || || ||
Db      398 LSQELAIPIYEE 408

```

Search completed: January 15, 2003, 11:20:14
Job time : 30.2857 secs

Sequence 545 AA;

Query Match	93.6%;	Score 44;	DB 13;	Length 545;
Best Local Similarity	81.8%;	Pred. No. 0.89;		
Matches	9;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;

```
QY      1 LSXELXIPYEE 11
        || || || || ||
Db      390 LSQELSIPIYEE 400
```

RESULT 15
2AF00304

ID AAE09304 standard; Protein; 552 AA.

AC AAEO9304;

DT 22-NOV-2001 (first entry)

Thermus aquaticus (Taq) DNA polymerase #1

KW *Thermus aquaticus*; Taq; DNA polymerase; polymerisation; PCR protocol, genetic engineering.

OS Thermus aquaticus.

PN W0200161015-A2.

PD 23-AUG-2001.
VY

PE 16-FEB-2001; 2001WO-EP01790.
YY

PR 17-FEB-2000; 2000US-0506153.
XX

PA (QIAG-) QIAGEN GMBH.
YY

PI Loeffert D, Missel A, Kang J;
YY

DR WPI; 2001-536571/59.
XY

PT Chimeric nucleic acid polymerase for nucleic acid polymerization and
 PT generation of nucleic acids, comprises two enzymatically active domains
 PT which are non-naturally associated .
 vv

PS Claim 12; Page 46-49; 93pp; English.
xy

CC The invention relates to a chimeric nucleic acid polymerase comprising
CC at least two enzymatically active domains which are non-naturally
CC associated. The polymerase is thermostable and possesses new or improved
CC catalytic properties compared to known nucleic acid polymerases. The
CC polymerase eliminates the need to specifically develop multi-enzyme
CC reaction mixtures, which are often difficult to optimise and expensive
CC to use. The polymerase facilitates rapid, efficient and accurate
CC generation of nucleic acid molecules, particularly in regard to PCR,
CC protocols. The polymerase is also useful for nucleic acid polymerisation
CC which is useful in genetic engineering techniques and molecular biology.
CC The present protein sequence is *Thermus aquaticus* (Taq) DNA polymerase.

552 AA;
Sequence

Query Match	93.6%;	Score 44;	DB 22;	Length 552;
Best Local Similarity	81.8%;	Pred. No. 0.9;		
Matches	9;	Conservative	0;	Mismatches 2;
			Indels	0;
			Gaps	0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 ; Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec

Title: US-09-823-649A-2
Perfect score: 47
Sequence: 1 LSXELXPYEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA: *
1: /cgn2.6/ptodata/1/iaa/5A.COMB.pep: *
2: /cgn2.6/ptodata/1/iaa/5B.COMB.pep: *
3: /cgn2.6/ptodata/1/iaa/6A.COMB.pep: *
4: /cgn2.6/ptodata/1/iaa/6B.COMB.pep: *
5: /cgn2.6/ptodata/1/iaa/PTCTUS.COMB.pep: *
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	93.6	528	2	US-08-484-956-90
2	44	93.6	528	2	US-09-757-653-90
3	44	93.6	528	4	US-08-520-946-90
4	44	93.6	553	4	US-09-096-399-2
5	44	93.6	554	1	US-08-021-623C-6
6	44	93.6	559	4	US-09-096-399-4
7	44	93.6	560	5	PCT-US95-14418-5
8	44	93.6	560	5	PCT-US95-15327-5
9	44	93.6	597	5	PCT-US95-14418-4
10	44	93.6	597	5	PCT-US95-15327-4
11	44	93.6	810	4	US-09-587-856-2
12	44	93.6	810	4	US-09-777-537-2
13	44	93.6	810	4	US-09-777-538-2
14	44	93.6	830	1	US-07-977-434-6
15	44	93.6	830	1	US-08-458-819-6
16	44	93.6	830	1	PCT-US91-07035-6
17	44	93.6	831	1	US-08-073-384C-5
18	44	93.6	831	1	US-08-254-359A-5
19	44	93.6	831	1	US-08-483-043-5
20	44	93.6	831	1	US-08-481-238-5
21	44	93.6	831	2	US-08-471-066B-5
22	44	93.6	831	2	US-08-484-956-5
23	44	93.6	831	2	US-08-757-653-5
24	44	93.6	831	2	US-08-599-491-5
25	44	93.6	831	2	US-08-756-386-5
26	44	93.6	831	2	US-08-823-516-5
27	44	93.6	831	3	US-08-662-853A-5

28	44	93.6	831	3	US-08-759-038-5	Sequence 5, Appl
29	44	93.6	831	3	US-08-758-314-5	Sequence 5, Appl
30	44	93.6	831	4	US-09-350-309-5	Sequence 5, Appl
31	44	93.6	831	4	US-08-520-946-5	Sequence 5, Appl
32	44	93.6	832	1	US-07-977-434-2	Sequence 2, Appl
33	44	93.6	832	1	US-08-156-020-2	Sequence 2, Appl
34	44	93.6	832	1	US-08-156-020-4	Sequence 4, Appl
35	44	93.6	832	1	US-08-156-020-6	Sequence 6, Appl
36	44	93.6	832	1	US-08-156-020-8	Sequence 8, Appl
37	44	93.6	832	1	US-08-156-020-10	Sequence 10, Appl
38	44	93.6	832	1	US-08-073-384C-4	Sequence 4, Appl
39	44	93.6	832	1	US-08-254-359A-4	Sequence 4, Appl
40	44	93.6	832	1	US-08-483-043-4	Sequence 4, Appl
41	44	93.6	832	1	US-08-458-819-2	Sequence 2, Appl
42	44	93.6	832	1	US-08-481-238-4	Sequence 4, Appl
43	44	93.6	832	2	US-08-471-066B-4	Sequence 4, Appl
44	44	93.6	832	2	US-08-484-956-4	Sequence 4, Appl
45	44	93.6	832	2	US-08-757-653-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-484-956-90
Sequence 90, Application US/08484956
Patent No. 5843654
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROW, MARY ANN D.
APPLICANT: OLSENBERG, MARY C.
APPLICANT: HERSLER, LAURA
TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL J., PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

Query Match 93.6%; Score 44; DB 4; length 553;
Best Local Similarity 81.8%; Pred. No. 0.36;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15327
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28003/31716
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15327-5

Query Match 93.6%; Score 44; DB 5; Length 560;
Best Local Similarity 81.8%; Pred. No. 0.37;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
DB 406 LSQELAIPIYEE 416

RESULT 9
PCT-US95-14418-4
Sequence 4, Application PC/TUS9514418
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA Encoding a Thermostable DNA Polymerase Enzyme
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14418
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28003/32330
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 597 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
PCT-US95-14418-4

Query Match 93.6%; Score 44; DB 5; Length 597;
Best Local Similarity 81.8%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
DB 443 LSQELAIPIYEE 453

RESULT 10
PCT-US95-15327-4
Sequence 4, Application PC/TUS9515327
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Biologically Active Fragments of
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/15327
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28003/31716
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 597 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-15327-4

Query Match 93.6%; Score 44; DB 5; Length 597;
Best Local Similarity 81.8%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
DB 443 LSQELAIPIYEE 453

RESULT 11
US-09-587-856-2
Sequence 2, Application US/09587856
Patent No. 6214557
GENERAL INFORMATION:
APPLICANT: Washington University
TITLE OF INVENTION: COLD SENSITIVE MUTANT DNA POLYMERASES
FILE REFERENCE: WSHU 2009
CURRENT APPLICATION NUMBER: US/09/587,856
CURRENT FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2

LENGTH: 810
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-587-856-2

Query Match 93.68; Score 44; DB 4; Length 810;
Best Local Similarity 81.88; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
DB 678 LSQELAIPIYEE 688

RESULT 12
US-09-777-537-2
Sequence 2, Application US/09777537
Patent No. 6316202
GENERAL INFORMATION:
APPLICANT: Barnes, Wayne M
APPLICANT: Kermekchlev, Milko B
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING COLD SENSITIVE MUTANT DNA
FILE REFERENCE: WSHU 2009.1
CURRENT APPLICATION NUMBER: US/09/777,537
CURRENT FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 09/587,856
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 810
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-777-537-2

Query Match 93.68; Score 44; DB 4; Length 810;
Best Local Similarity 81.88; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
DB 678 LSQELAIPIYEE 688

RESULT 13
US-09-777-538-2
Sequence 2, Application US/09777538
Patent No. 633159
GENERAL INFORMATION:
APPLICANT: Barnes, Wayne M
APPLICANT: Kermekchlev, Milko B
TITLE OF INVENTION: COLD SENSITIVE MUTANT DNA POLYMERASES AND METHODS OF
FILE REFERENCE: WSHU 2009.2
CURRENT APPLICATION NUMBER: US/09/777,538
CURRENT FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 09/587,856
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 810
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-777-538-2

Query Match 93.68; Score 44; DB 4; Length 810;
Best Local Similarity 81.88; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
DB 678 LSQELAIPIYEE 688

DB 678 LSQELAIPIYEE 688

RESULT 14
US-07-977-434-6
Sequence 6, Application US/07977434
Patent No. 546591
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: Wordperfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseer
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 546591 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-977-434-6

Query Match 93.6%; Score 44; DB 1; Length 830;
Best Local Similarity 81.8%; Pred. No. 0.56;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
||| |||||
Db 676 LSQELSIPIYEE 686

RESULT 15
US-08-458-819-6
Sequence 6, Application US/08458819
Patent No. 5795762
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: wordperfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,819
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,434
FILING DATE: 23-FEB-1993
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471

FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseif
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 5795762 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-819-6

Query Match 93.6%; Score 44; DB 1; Length 830;
Best Local Similarity 81.8%; Pred. No. 0.56;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11
||| |||||
Db 676 LSQELSIPIYEE 686

Search completed: January 15, 2003, 12:38:22
Job time: 11.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:20:34 ; Search time 10.8571 Seconds
(without alignments)
20.140 Million cell updates/sec

Title: US-09-823-649A-2
Perfect score: 47
Sequence: 1 LSEXLAIPYEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US06_NMW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCU_NMW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NMW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NMW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCU05_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NMW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NMW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NMW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	93.6	11	10	US-09-823-649A-3
2	44	93.6	11	10	US-09-823-649A-4
3	44	93.6	11	10	US-09-823-649A-8
4	44	93.6	11	10	US-09-823-649A-9
5	44	93.6	11	10	US-09-823-649A-10
6	44	93.6	11	10	US-09-823-649A-11
7	44	93.6	11	10	US-09-823-649A-12
8	44	93.6	11	10	US-09-823-649A-13
9	44	93.6	11	10	US-09-823-649A-14
10	44	93.6	831	9	US-10-033-297-5
11	44	93.6	831	9	US-10-081-806-5
12	44	93.6	832	9	US-10-033-297-4
13	44	93.6	832	9	US-09-872-834-2
14	44	93.6	832	9	US-10-071-805-1
15	44	93.6	832	9	US-10-081-806-4
16	44	93.6	833	9	US-10-033-297-8
17	44	93.6	833	9	US-10-033-297-66
18	44	93.6	833	9	US-10-033-297-69
19	44	93.6	833	9	US-10-033-297-71

20	44	93.6	833	9	US-10-081-806-8	Sequence 8, Appl1
21	44	93.6	834	9	US-10-033-297-6	Sequence 6, Appl1
22	44	93.6	834	9	US-10-081-806-6	Sequence 6, Appl1
23	44	93.6	836	10	US-09-777-430A-8	Sequence 8, Appl1
24	44	93.6	836	10	US-09-777-430A-11	Sequence 11, Appl1
25	44	93.6	842	10	US-09-777-430A-15	Sequence 15, Appl1
26	44	93.6	842	10	US-09-777-430A-20	Sequence 20, Appl1
27	44	93.6	842	10	US-09-777-430A-23	Sequence 23, Appl1
28	44	93.6	842	10	US-09-777-430A-26	Sequence 26, Appl1
29	43	91.5	11	10	US-09-823-649A-2	Sequence 2, Appl1
30	39	83.0	560	9	US-10-071-505-2	Sequence 2, Appl1
31	39	83.0	830	9	US-10-071-505-3	Sequence 3, Appl1
32	38	80.9	11	12	US-10-053-632-3	Sequence 3, Appl1
33	38	80.9	11	12	US-10-053-632-14	Sequence 14, Appl1
34	38	80.9	11	12	US-10-052-417-3	Sequence 3, Appl1
35	38	80.9	11	12	US-10-052-417-14	Sequence 14, Appl1
36	37	78.7	11	12	US-10-053-632-2	Sequence 2, Appl1
37	37	78.7	11	12	US-10-053-632-13	Sequence 13, Appl1
38	37	78.7	11	12	US-10-052-417-2	Sequence 2, Appl1
39	37	78.7	11	12	US-10-052-417-13	Sequence 13, Appl1
40	33	70.2	126	10	US-09-734-017A-58	Sequence 58, Appl1
41	33	70.2	499	10	US-09-735-169A-2	Sequence 2, Appl1
42	33	70.2	499	10	US-09-735-171A-2	Sequence 2, Appl1
43	32	68.1	194	10	US-09-925-302-585	Sequence 585, App
44	31	66.0	375	12	US-10-062-254-373	Sequence 373, App
45	30	63.8	215	10	US-09-854-845-19	Sequence 19, Appl1

ALIGNMENTS

```
RESULT 1
US-09-823-649A-3
; Sequence 3, Application US/09823649A
; Patent No. US20020012970A1
;
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elfstrom, Carla
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrenner, Nancy
;
TITLE OR INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
;
FILE REFERENCE: RPAL1006
;
CURRENT APPLICATION NUMBER: US/09/823,649A
;
PRIORITY APPLICATION NUMBER: 2001-04-30
;
PRIORITY FILING DATE: 2001-04-18
;
NUMBER OF SEQ ID NOS: 21
;
SOFTWARE: PatentIn version 3.0
;
SEQ ID NO 3
;
LENGTH: 11
;
TYPE: PRT
;
ORGANISM: Artificial Sequence
;
FEATURE:
;
OTHER INFORMATION: sequence motif
US-09-823-649A-3
;
Query Match          93.6%; Score 44; DB 10; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0021;
Matches 9; Conservative 0; Mismatches 2; Indels 0;
;
RESULT 2
US-09-823-649A-4
; Sequence 4, Application US/09823649A
; Patent No. US20020012970A1
;
GENERAL INFORMATION:
```

```
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
FILE REFERENCE: RPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: sequence motif
NAME/KEY: VARIANT
LOCATION: (3)..(3)
OTHER INFORMATION: X is Q or G
US-09-823-649A-4
```

```
Query Match          93.6%; Score 44; DB 10; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0021;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 LSXELIPIYEE 11
   ||||| |||||
DB 1 LSGELAIPIYEE 11
```

```
RESULT 3
US-09-823-649A-8
Sequence 8, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
FILE REFERENCE: RPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 11
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-823-649A-8
```

```
Query Match          93.6%; Score 44; DB 10; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0021;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 LSXELIPIYEE 11
   ||||| |||||
DB 1 LSGELAIPIYEE 11
```

```
RESULT 4
US-09-823-649A-9
Sequence 9, Application US/09823649A
```

```
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
FILE REFERENCE: RPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 11
TYPE: PRT
ORGANISM: Thermus flavus
US-09-823-649A-9
```

```
Query Match          93.6%; Score 44; DB 10; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0021;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 LSXELIPIYEE 11
   ||||| |||||
DB 1 LSGELAIPIYEE 11
```

```
RESULT 5
US-09-823-649A-10
Sequence 10, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
FILE REFERENCE: RPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 11
TYPE: PRT
ORGANISM: Thermus thermophilus
US-09-823-649A-10
```

```
Query Match          93.6%; Score 44; DB 10; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0021;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 LSXELIPIYEE 11
   ||||| |||||
DB 1 LSGELAIPIYEE 11
```

```
RESULT 6
US-09-823-649A-11
Sequence 11, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
```

```

; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus sp. Z05
US-09-823-649A-11

```

```

Query Match
Best Local Similarity 93.6%; Score 44; DB 10; Length 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy 1 LSXELXPYEE 11
Db 1 LSQELAIPTYEE 11

```

```

RESULT 7
; Sequence 12, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus sp. sps17
US-09-823-649A-12

```

```

Query Match
Best Local Similarity 93.6%; Score 44; DB 10; Length 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy 1 LSXELXPYEE 11
Db 1 LSQELAIPTYEE 11

```

```

RESULT 8
; Sequence 13, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell

```

```

; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus caldophilus
US-09-823-649A-13

```

```

Query Match
Best Local Similarity 93.6%; Score 44; DB 10; Length 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy 1 LSXELXPYEE 11
Db 1 LSQELAIPTYEE 11

```

```

RESULT 9
; Sequence 14, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus 11111ormis
US-09-823-649A-14

```

```

Query Match
Best Local Similarity 93.6%; Score 44; DB 10; Length 11;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
Qy 1 LSXELXPYEE 11
Db 1 LSQELAIPTYEE 11

```

```

RESULT 10
; Sequence 5, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brown, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163

```


;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medlen & Carroll, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94104
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/033,297
;; FILING DATE: 12-NOV-2002
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/350,597
;; FILING DATE: 09-JUL-1999
;; APPLICATION NUMBER: US/08/823,516
;; FILING DATE: 24-MAR-1997
;; APPLICATION NUMBER: PCT/US97/01072
;; FILING DATE: 21-JAN-1997
;; APPLICATION NUMBER: US 08/759,038
;; FILING DATE: 02-DEC-1996
;; APPLICATION NUMBER: US 08/758,314
;; FILING DATE: 02-DEC-1996
;; APPLICATION NUMBER: US 08/756,386
;; FILING DATE: 29-NOV-1996
;; APPLICATION NUMBER: US 08/682,853
;; FILING DATE: 12-JUL-1996
;; APPLICATION NUMBER: US 08/599,491
;; FILING DATE: 24-JAN-1996
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ingolia, Diane F.
;; REGISTRATION NUMBER: 40,027
;; REFERENCE/DOCKET NUMBER: F0RS-02736
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;;
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 831 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-033-297-5
;
Query Match 93.6%; Score 44; DB 9; Length 831;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
QY 1 LSXELXPYEE 11
DB 677 LSSELSTPYEE 687
;
RESULT 11
US-10-081-806-5
;; Sequence 5, Application US/10081806
;; Publication No. US20020197623A1
;; GENERAL INFORMATION:
;; APPLICANT: Prudent, James R.
;; Hall, Jeff G.
;; Lyamichev, Victor I.
;; TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medlen & Carroll, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco

;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94104
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/081,806
;; FILING DATE: 22-Feb-2002
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/756,386
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: US 08/682,853
;; FILING DATE: 12-JUL-1996
;; APPLICATION NUMBER: US 08/599,491
;; FILING DATE: 24-JAN-1996
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ingolia, Diane F.
;; REGISTRATION NUMBER: 40,027
;; REFERENCE/DOCKET NUMBER: F0RS-02564
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;;
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 831 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-081-806-5
;
Query Match 93.6%; Score 44; DB 9; Length 831;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
QY 1 LSXELXPYEE 11
DB 677 LSSELSTPYEE 687
;
RESULT 12
US-10-033-297-4
;; Sequence 4, Application US/10033297
;; Publication No. US20020187486A1
;; GENERAL INFORMATION:
;; APPLICANT: Hall, Jeff G.
;; Lyamichev, Victor I.
;; Mast, Andrea L.
;; Brow, Mary Ann D.
;; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;; Sequential Invasive Cleavages
;; NUMBER OF SEQUENCES: 163
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medlen & Carroll, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94104
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/033,297
;; FILING DATE: 12-NOV-2002
;; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-Jul-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,366
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020187486A1 Relevant
TOPOLOGY: No. US20020187486A1 Relevant
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-033-297-4
Query Match 93.6%; Score 44; DB 9; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 LSXELXIPYE 11
DB 678 LSQELAIPIYE 688
RESULT 13
US-09-972-834-2
Sequence 2, Application US/09972834
Publication No. US2002019263A1
GENERAL INFORMATION:
APPLICANT: Loeb, Lawrence A.
Hood, Leroy
Suzuki, Motohshi
TITLE OF INVENTION: Thermostable Polymerases Having Altered
Fidelity and Methods of Identifying and Using Same
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL & FLORES LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,834
FILING DATE: 04-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/978,806
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-0W 2873
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-972-834-2
Query Match 93.6%; Score 44; DB 9; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 LSXELXIPYE 11
DB 678 LSQELAIPIYE 688
RESULT 14
US-10-071-505-1
Sequence 1, Application US/10071505
Publication No. US2002019721A1
GENERAL INFORMATION:
APPLICANT: Davis, Maria
APPLICANT: Nelson, John
APPLICANT: Kumar, Shiv
APPLICANT: Flinn, Patrick J.
APPLICANT: Nampalli, Salyam
TITLE OF INVENTION: PAC DNA Polymerase Having an Amino Acid Substitution at
TITLE OF INVENTION: E681 and Homologs thereof Exhibiting Improved Salt
Tolerance
FILE REFERENCE: PB9944
CURRENT APPLICATION NUMBER: US/10/071,505
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/148,012
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent In Vit. 2.1
SEQ ID NO: 1
LENGTH: 832
TYPE: PRT
ORGANISM: Thermus aquaticus
US-10-071-505-1
Query Match 93.6%; Score 44; DB 9; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 LSXELXIPYE 11
DB 678 LSQELAIPIYE 688
RESULT 15
US-10-081-806-4
Sequence 4, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-081-806-4
Query Match 93.64; Score 44; DB 9; Length 832;
Best Local Similarity 81.88; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 LSXELXIPYEE 11
||| |||||
Db 678 LSQELAIYEE 688
Search completed: January 15, 2003, 12:39:50
Job time: 10.8571 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:15:24 ; Search time 12.1429 seconds
(without alignments)
87.086 Million cell updates/sec

Title: US-09-823-649A-2

Perfect score: 47

Sequence: LSXELXIPYEE 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	44	93.6	831	2	DNA-directed DNA p
2	44	93.6	831	2	DNA-directed DNA p
3	44	93.6	831	2	DNA-directed DNA p
4	38	80.9	229	2	hypothetical prote
5	38	80.9	850	2	conserved hypotet
6	37	78.7	457	2	alpha-galactosidas
7	37	78.7	474	2	alpha-galactosidas
8	36	76.6	1451	2	B86286
9	35	74.5	149	1	W6WL31
10	35	74.5	327	2	JC7337
11	35	74.5	634	2	C70133
12	34	72.3	881	2	T30452
13	34	72.3	738	2	T27289
14	34	72.3	921	2	A40597
15	33	70.2	233	2	AB1440
16	33	70.2	233	2	AI1082
17	33	70.2	326	2	S25321
18	33	70.2	390	2	S75715
19	33	70.2	433	2	G64312
20	33	70.2	597	2	T40331
21	33	70.2	692	2	G90284
22	33	70.2	1136	2	JQ0494
23	33	70.2	4717	2	T41581
24	32	68.1	1332	2	G84529
25	32	68.1	153	2	AC1431
26	32	68.1	153	2	AI1804
27	32	68.1	186	2	H98220
28	32	68.1	186	2	AH3065
29	32	68.1	197	2	H64415

ALIGNMENTS

30	32	68.1	228	2	A53028
31	32	68.1	402	2	G72516
32	32	68.1	417	2	E75148
33	32	68.1	469	2	G69267
34	32	68.1	753	1	KXMS01
35	32	68.1	1062	2	B26350
36	32	68.1	1234	2	B36186
37	32	68.1	1493	2	E70435
38	32	68.1	1495	2	S60255
39	32	68.1	1890	2	T04556
40	31	66.0	144	2	AF1445
41	31	66.0	144	2	AF1362
42	31	66.0	153	2	G82778
43	31	66.0	161	2	G83727
44	31	66.0	162	1	DEBSPE
45	31	66.0	193	2	D97157

RESULT 1

S26675

DNA-directed DNA polymerase (EC 2.7.7.7) I - Thermus aquaticus

C:Species: Thermus aquaticus

C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999

C:Accession: S26675; S24929

R:Akmetzjanov, A.A.; Vakhitov, V.A.

Nucleic Acids Res. 20, 5839, 1992

A:Title: Molecular cloning and nucleotide sequence of the DNA polymerase gene from Th

A:Reference number: S26675; MUID:93087201; PMID:1454544

A:Accession: S26675

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-831 <AKH>

A:Cross-references: EMBL:X66105; NID:948165; PIDN:CAA46900.1; PID:948166

A>Note: The source is designated as Thermus flavus

C:Genetics:

A:Gene: pol I

A:Superfamily: DNA-directed DNA polymerase I

C:Keywords: DNA binding; nucleotidyltransferase

Query Match

93.6%: Score 44; DB 2; Length 831;

Best Local Similarity 81.8%: Pred. No. 0.46;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11

DB 677 LSXELXIPYEE 687

RESULT 2

JX0359

DNA-directed DNA polymerase (EC 2.7.7.7) - Thermus aquaticus

C:Species: Thermus aquaticus

C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 07-May-1999

C:Accession: JX0359

R:Rishino, Y.; Ueno, T.; Miyagi, M.; Demori, T.; Imanura, M.; Tsunazawa, S.; Kato, I.

J. Biochem. 116, 1019-1024, 1994

A:Title: Overproduction of Thermus aquaticus DNA polymerase and its structural analys

A:Reference number: JX0359; MUID:95204371; PMID:7896728

A:Accession: JX0359

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-831 <TSH>

C:Superfamily: DNA-directed DNA polymerase I

C:Keywords: nucleotidyltransferase

Query Match

93.6%: Score 44; DB 2; Length 831;

Best Local Similarity 81.8%: Pred. No. 0.46;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELXIPYEE 11

Db 678 LSQELAIPIYEE 688

RESULT 3

A33530
DNA-directed DNA polymerase (EC 2.7.7.7) I - *Thermus aquaticus*
C:Species: *Thermus aquaticus*
C>Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 18-Jun-1999
C:Accession: A33530
R:Lewyer, F.C.; Stoffel, S.; Saito, R.K.; Myambo, K.; Drummond, R.; Gelfand, D.H.
J. Biol. Chem. 264, 6427-6437, 1989
A:Title: Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase I
A:Reference number: A33530; MUID:89197950; PMID:2649500
A:Accession: A33530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-832 <LAN>
A:Cross-references: GB:J04639; NID:q155128; PIDN:AAA27507.1; PID:q155129; GB:X26480
C:Superfamily: DNA-directed DNA polymerase I
C:Keywords: DNA binding; nucleotidyltransferase

Query Match 93.6%; Score 44; DB 2; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.46;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELAIPIYEE 11
||| |||||
Db 678 LSQELAIPIYEE 688

RESULT 4

S46696
Hypothetical protein YHR071W - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein H8025.1
C:Species: *Saccharomyces cerevisiae*
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 29-Oct-1999
C:Accession: S46696
R:Latrelle, P.
Submitted to the EMBL Data Library, May 1994
A:Description: The sequence of *S. cerevisiae* cosmid 8025.
A:Reference number: S46696
A:Accession: S46696
A:Molecule type: DNA
A:Residues: 1-229 <LAN>
A:Cross-references: EMBL:U00061; NID:q487943; PIDN:AB68375.1; PID:q487944; MIPS:YHR071W
C:Genetics:
A:Gene: SGD:PCLS
A:Cross-references: SGD:S000113; MIPS:YHR071W
A:Map position: 8R

Query Match

80.9%; Score 38; DB 2; Length 229;
Best Local Similarity 63.6%; Pred. No. 1.9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELAIPIYEE 11
||| |||||
Db 174 LNYELAIPIYDE 184

RESULT 5

G70332
Conserved hypothetical protein aq_367 - *Aquifex aeolicus*
C:Species: *Aquifex aeolicus*
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: G70332
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: G70332
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-850 <AGF>
A:Cross-references: GB:AE000686; NID:g2983038; PIDN:AA06656.1; PID:g2983047; GB:AE00
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_367

Query Match 80.9%; Score 38; DB 2; Length 850;
Best Local Similarity 54.5%; Pred. No. 8.5;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELAIPIYEE 11
||| |||||
Db 227 ISREKDVPIYEE 237

RESULT 6

AE3130
alpha-galactosidase melA [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C:Species: *Agrobacterium tumefaciens*
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 17-May-2002
C:Accession: AE3130
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.R.; Chen, Y.; Woerage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kalyavin, T.; Levy, R.; Li, M.; McCl ; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.
A:Title: The genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AE3130
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <KUR>
A:Cross-references: GB:AE008683; PIDN:AA14545.1; PID:q17743164; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: melA
A:Map position: linear chromosome
C:Superfamily: melibiose-specific alpha-galactosidase

Query Match 78.7%; Score 37; DB 2; Length 457;
Best Local Similarity 63.6%; Pred. No. 6.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELAIPIYEE 11
||| |||||
Db 183 LAMHDIPIYEE 193

RESULT 7

D98157
alpha-galactosidase (melibiase) [imported] - *Agrobacterium tumefaciens* (strain C58, C
C:Species: *Agrobacterium tumefaciens*
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002
C:Accession: D98157
R:Gochner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*
A:Reference number: A97359; PMID:11743194
A:Accession: D98157
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-474 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK86782.1; PID:q15156531; GSPDB:GN00170
C:Genetics:
A:Gene: AgR_L436
A:Map position: linear chromosome
C:Superfamily: melibiose-specific alpha-galactosidase

Query Match 78.7%; Score 37; DB 2; Length 474;
Best Local Similarity 63.6%; Pred. No. 7.1;

RT adrenergic receptor. Identification of a unique receptor subtype.";
 RL J. Biol. Chem. 266:10470-10478(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Saitoh M., Imai A., Shimomura H.;
 RT Cloning of rat alpha-2-B-adrenergic receptor gene and expression in
 RL rat submandibular gland.";
 RL Shigenaka 80:317-326(1992).
 CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIANE THE CATECHOLAMINE-
 CC INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS.
 CC SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M58316; AAA40634.1; -;
 DR EMBL: X57659; CAA40861.1; -;
 DR EMBL: M62371; AAA42033.1; -;
 DR EMBL: D00819; BAA00700.1; -;
 DR PIR: A37869; A37869.
 DR PIR: S13023; S13023.
 DR PIR: A40392; A40392.
 DR HSP: P29274; 1XMH.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR000237; GPCR_Rhodopsin.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Phosphorylation.
 FT DOMAIN 1 51
 FT TRANSMEM 52 76
 FT DOMAIN 77 88
 FT TRANSMEM 89 114
 FT DOMAIN 115 124
 FT TRANSMEM 125 147
 FT DOMAIN 148 168
 FT TRANSMEM 169 191
 FT DOMAIN 192 207
 FT TRANSMEM 208 231
 FT DOMAIN 232 379
 FT TRANSMEM 380 403
 FT DOMAIN 404 416
 FT TRANSMEM 417 437
 FT DOMAIN 438 458
 FT TRANSMEM 459 484
 FT CARBOHYD 19 19
 FT CARBOHYD 33 33
 FT DISULFID 124 202
 FT CONFLICT 24 24
 FT CONFLICT 40 40
 FT CONFLICT 69 69
 FT CONFLICT 155 155
 FT CONFLICT 245 245
 FT CONFLICT 252 252
 FT CONFLICT 275 275
 FT CONFLICT 298 298
 SQ SEQUENCE 458 AA; 49864 MW; 6846C2AF6333BF CRC64;
 Query Match 44.4%; Score 8; DB 1; Length 458;
 Best Local Similarity 18.2%; Pred. No. 9.1e-07;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

RESULT 11
 RPN4_YEAST
 ID RPN4_YEAST STANDARD: PRT; 531 AA.
 AC Q03465;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 26S Proteasome regulatory subunit RPN4 (Nuclear protein SON1) (UB
 DE fusion degradation protein 5).
 GN RPN4 OR SON1 OR UFD5 OR YDL020C OR D2840.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Eungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 RX MEDLINE:9329218; PubMed-8514125;
 RA Nelson M.K., Kurihara T., Silver P.A.;
 RT "Extragenic suppressors of mutations in the cytoplasmic C terminus of
 RT SPC63 define five genes in Saccharomyces cerevisiae.";
 RL Genetics 134:159-173(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-N303;
 RC MEDLINE-9329218; PubMed-8514125;
 RA Nelson M.K., Kurihara T., Silver P.A.;
 RT "Extragenic suppressors of mutations in the cytoplasmic C terminus of
 RT SPC63 define five genes in Saccharomyces cerevisiae.";
 RL Genetics 134:159-173(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Andre B., Vissers S., Urrestarazu L.;
 RC STRAIN-S288C;
 RA Submitted (F85-1995) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RN CHARACTERIZATION.
 RC STRAIN-S288C;
 RX MEDLINE-95340540; PubMed-7615550;
 RA Johnson E.S., Ma J.C.M., Ota I.M., Varshavsky A.;
 RT "A proteolytic pathway that recognizes ubiquitin as a degradation
 RT signal.";
 RL J. Biol. Chem. 270:17442-17456(1995).
 RN [4]
 RN CHARACTERIZATION.
 RX MEDLINE-98171302; PubMed-9512348;
 RA Fujimuro M., Tanaka K., Yokosawa H., Toh-E A.;
 RT "Son1p is a component of the 26S proteasome of the yeast
 RT Saccharomyces cerevisiae.";
 RL FEBS Lett. 423:149-154(1998).
 CC -1- FUNCTION: MAY PLAY A ROLE IN NUCLEAR INTEGRITY, IS REQUIRED FOR
 CC NORMAL GROWTH AT LOW TEMPERATURES. SON1 MUTANTS GROW SLOWLY AT LOW
 CC TEMPERATURES AND SHOW PARTIAL MISLOCALIZATION OF NUCLEAR ANTIGENS.
 CC PROBABLY INITIATES WITH SEC63. COMPONENT OF 26S PROTEASOME.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: I00928; AAA35067.1; -;
 DR EMBL: Z48432; CAA83339.1; -;
 DR EMBL: Z74068; CAA98579.1; -;
 DR PIR: S30806; S30806.
 DR PIR: S41986; S41986.
 DR TRANSFAC: T04539; -;
 DR SCD: S0002178; RPN4.
 DR InterPro: IPR000822; ZnF_C2H2.
 DR Pfam: PF00096; ZnF_C2H2_2.
 DR SMART: SM00355; ZnF_C2H2_1.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
 KW Proteasome; Nuclear protein.
 FT DOMAIN 211 229
 FT DOMAIN 300 315
 FT DOMAIN 382 398
 ASP/GLU-RICH (ACIDIC).
 ASP/GLU-RICH (ACIDIC).
 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SQ SEQUENCE 531 AA; 60152 MW; 4316281AC09FE97F CRC64;
 Query Match 44.4%; Score 8; DB 1; Length 531;
 Best Local Similarity 18.2%; Pred. No. 9.8e-07;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 LXXXXXXXE 11
 DB 94 LTTTSTTSE 104
 RESULT 12
 ID VG12_BPP2 STANDARD; PRT; 854 AA.
 AC P20345;
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Pre-neck appendage protein (late protein GP12).
 GN 12.
 OS Bacteriophage phi-29.
 CC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 CC phi-29-like viruses.
 OX NCBI_TaxID=10756;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87106857; PubMed=3803926;
 RA Vleck C.; Paces V.;
 RT "Nucleotide sequence of the late region of Bacillus phage phi 29
 RT completes the 19,285-bp sequence of phi 29 genome. Comparison with
 RT the homologous sequence of phage PZA.";
 RC Gene 46:215-225(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M14782; AAA2285.1; -.
 DR PIR; G25816; WMBP22.
 KW Late protein.
 SQ SEQUENCE 854 AA; 92102 MW; 373D4F4275D37EED CRC64;
 Query Match 44.4%; Score 8; DB 1; Length 854;
 Best Local Similarity 18.2%; Pred. No. 1.3e-06;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 LXXXXXXXE 11
 DB 587 LTSSSKTSE 597
 RESULT 13
 ID VG12_BPP2A STANDARD; PRT; 854 AA.
 AC P07537;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Pre-neck appendage protein (late protein GP12).
 GN 12.
 OS Bacteriophage PZA.
 CC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 CC phi-29-like viruses.
 OX NCBI_TaxID=10757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87031573; PubMed=3095188;
 RA Paces V.; Vleck C.; Urbanek P.;
 RT "Nucleotide sequence of the late region of Bacillus subtilis phage

RT PZA, a close relative of phi 29.";
 RL Gene 44:107-114(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M1813; AA08489.1; -.
 DR PIR; G24831; WMBP12.
 KW Late protein.
 SQ SEQUENCE 854 AA; 92072 MW; 3FB2P23652DB8E3F CRC64;
 Query Match 44.4%; Score 8; DB 1; Length 854;
 Best Local Similarity 18.2%; Pred. No. 1.3e-06;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 LXXXXXXXE 11
 DB 587 LTSSSKTSE 597
 RESULT 14
 ID YMG8_YEAST STANDARD; PRT; 953 AA.
 AC Q03516;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 107.7 kDa protein in TSP3-IPR2 intergenic region.
 GN YMR266W OR YMR156.08.
 OS Saccharomyces cerevisiae (baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AH972;
 RA Lye G.; Churcher C.M.; Barrell B.G.; Rajandream M.A.;
 RT Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BRLONGS TO THE YL005C/YMR266W/YOL084W FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z49260; CA089249.1; -.
 DR SGD; S0004879; YMR266W.
 DR InterPro; IPR003864; DUF221.
 DR Pfam; PF02714; DUF221; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 103 123 POTENTIAL.
 FT TRANSMEM 148 168 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT TRANSMEM 392 412 POTENTIAL.
 FT TRANSMEM 435 455 POTENTIAL.
 FT TRANSMEM 481 501 POTENTIAL.
 FT TRANSMEM 540 560 POTENTIAL.
 FT TRANSMEM 575 595 POTENTIAL.
 FT TRANSMEM 599 619 POTENTIAL.
 FT TRANSMEM 642 662 POTENTIAL.
 FT TRANSMEM 666 686 POTENTIAL.
 SQ SEQUENCE 953 AA; 107672 MW; F6F7BA2B3F3AD44B CRC64;

Query Match 44.4%; Score 8; DB 1; Length 953;
 Best Local Similarity 18.2%; Pred. No. 1.3e-06;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXX 11
 DB 496 LSSAATVTE 506

RESULT 15

EMBL_MOUSE STANDARD; PRT; 988 AA.
 AC 070318;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Band 4.1-like protein 2 (Generally expressed protein 4.1) (4.16).
 GN EPR41L2 OR EPR4.1L2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98198473; PubMed=9531554;
 RA Walensky L.D., Gascard P., Fields M.E., Blackshaw S., Conboy J.G.,
 R1 Mohandas N., Snyder S.H.;
 RT "The 13-kD FK506 binding protein, FKBP13, interacts with a novel
 RL homologue of the erythrocyte membrane cytoskeletal protein 4.1.";
 RL J. Cell Biol. 141:143-153(1998).
 CC -1- SUBUNIT: The CYP domain interacts with FKBP-13.
 CC -1- TISSUE SPECIFICITY: Widely expressed.
 CC -1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; AF044312; AAC40083.1; -.
 DR MGD; MGI:103009; Epp4.112.
 DR InterPro; IPR000299; Band.4.1.
 DR Pfam; PF00373; Band.41; 1.
 DR PRINTS; PR00935; BAND41.
 DR SMART; SM00295; B41; 1.
 DR PROSITE; PS00660; BAND_41_1; 1.
 DR PROSITE; PS00661; BAND_41_2; 1.
 DR PROSITE; PS00657; BAND_41_3; 1.
 KW Structural protein; Cytoskeleton.
 FT DOMAIN 208 422 BAND 4.1-LIKE.
 FT DOMAIN 489 651 HYDROPHILIC.
 FT DOMAIN 652 837 SPECTRIN--ACTIN-BINDING.
 FT DOMAIN 838 988 CARBOXYL-TERMINAL (CTD).
 SQ SEQUENCE 988 AA; 109833 MW; B0367A16C5A2EC05 CRC64;

Query Match 44.4%; Score 8; DB 1; Length 988;
 Best Local Similarity 18.2%; Pred. No. 1.4e-06;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXX 11
 DB 720 LSSSSSSSSSE 730

Search completed: January 15, 2003, 11:21:14
 Job time : 13.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: January 15, 2003, 10:29:39 ; Search time 25.5714 Seconds

(without alignments)
88.635 Million cell updates/sec

Title: US-09-823-649A-1

Perfect score: 18
Sequence: 1 LXXXXXXXXXXE 11

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	50.0	101	12 Q9J1A0	Q9J1A0 pseudorabies
2	9	50.0	274	5 Q9VNY3	Q9VNY3 drosophila
3	9	50.0	303	10 Q81807	Q81807 arabidopsis
4	9	50.0	321	10 Q9ZUS8	Q9ZUS8 arabidopsis
5	9	50.0	326	10 Q9PH69	Q9PH69 arabidopsis
6	9	50.0	466	4 Q9H0C7	Q9H0C7 homo sapien
7	9	50.0	639	4 Q96T25	Q96T25 homo sapien
8	9	50.0	761	5 Q9BHY3	Q9BHY3 leishmania
9	9	50.0	803	5 Q9VYV9	Q9VYV9 drosophila
10	9	50.0	1133	4 Q60288	Q60288 homo sapien
11	9	50.0	2207	5 Q9UOV2	Q9UOV2 leishmania
12	9	50.0	2768	5 Q9VC00	Q9VC00 drosophila
13	8	44.4	33	4 Q9DMS8	Q9DMS8 homo sapien
14	8	44.4	56	10 Q9JL43	Q9JL43 trepouxia j
15	8	44.4	60	5 Q9NMB8	Q9NMB8 leishmania
16	8	44.4	78	3 Q8WZM6	Q8WZM6 neurospora

17	8	44.4	86	15 Q87745	Q87745 chimpanzee
18	8	44.4	87	9 Q9G0C8	Q9G0C8 lactococcus
19	8	44.4	91	10 Q942X0	Q942X0 oryza sativ
20	8	44.4	99	5 Q9VPS9	Q9VPS9 drosophila
21	8	44.4	107	5 Q9VPS8	Q9VPS8 drosophila
22	8	44.4	119	15 Q87741	Q87741 chimpanzee
23	8	44.4	140	11 Q64145	Q64145 ratius sp.
24	8	44.4	157	4 Q8TD04	Q8TD04 homo sapien
25	8	44.4	160	5 Q9GP85	Q9GP85 cryptospori
26	8	44.4	169	16 Q9RMU6	Q9RMU6 delnococtus
27	8	44.4	178	4 Q9HBF2	Q9HBF2 homo sapien
28	8	44.4	185	5 Q9W3S3	Q9W3S3 drosophila
29	8	44.4	189	16 Q8UE55	Q8UE55 agrobacteri
30	8	44.4	192	4 Q9UHA6	Q9UHA6 homo sapien
31	8	44.4	197	4 Q8TD05	Q8TD05 homo sapien
32	8	44.4	204	11 Q9RPM8	Q9RPM8 mus musculu
33	8	44.4	207	16 Q9AK65	Q9AK65 caulobacter
34	8	44.4	213	10 Q8SB85	Q8SB85 oryza sativ
35	8	44.4	215	4 Q9HIC6	Q9HIC6 homo sapien
36	8	44.4	218	4 Q9NZH6	Q9NZH6 homo sapien
37	8	44.4	218	4 Q9HBF3	Q9HBF3 homo sapien
38	8	44.4	233	5 Q9VVM9	Q9VVM9 drosophila
39	8	44.4	233	5 Q8SZP1	Q8SZP1 drosophila
40	8	44.4	238	10 Q941H2	Q941H2 nanochlori
41	8	44.4	259	11 Q9R1I2	Q9R1I2 microtus oc
42	8	44.4	261	5 Q9NBY3	Q9NBY3 graneledone
43	8	44.4	288	5 Q9VT16	Q9VT16 drosophila
44	8	44.4	298	2 Q9AMM4	Q9AMM4 bradyrhizob
45	8	44.4	315	11 Q9JK65	Q9JK65 ratius norv

ALIGNMENTS

Q9J1A0	PRELIMINARY;	PRT;	101 AA.
AC Q9J1A0			
DT 01-OCT-2000 (TREMBLrel. 15, Created)			
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE IRI80 (Fragment).			
GN IRI80.			
OS Pseudorabies virus.			
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC Alphaherpesvirinae; Varicellovirinae.			
OC NCBI_TaxID=10345;			
PN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=PSEUDORABIES VIRUS STRAIN RA;			
RA Fang L.R., Xiao S.B., Xu J.X., Hong W.Z., Chen H.C.;			
RT "Cloning and expression of IRI80 gene of pseudorabies virus Ea strain."			
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
DR EMBL: AF218842; AAF43385.1; -			
FT NON_TER			
FT 101			
SQ SEQUENCE 101 AA; 10187 MW; 6E45D3887C53CA94 CRC64;			
Query Match	50.0%;	Score 9;	DB 12;
Best Local Similarity	18.2%;	Pred. No. 5.7e-14;	Length 101;
Matches 2;	Conservative	0;	Mismatches 9;
			Indels 0;
			Gaps 0;
QY 1 LXXXXXXXXXXE 11			
DB 19 LAAAAAAAAAAE 29			
RESULT 2			
Q9VNY3	PRELIMINARY;	PRT;	274 AA.
ID Q9VNY3			
AC Q9VNY3			
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			

```

DT 01-MAY-2000 (Tremblrel. 13, last annotation update)
DE CG14564 protein.
GN CG14564.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phytotidae; Drosophilidae; Drosophila.
XX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE-20196006; PubMed-10731132.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Plamkosh C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkeva D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dun P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.D., Mei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacheco J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith V.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Ventler J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003595; AAF51780.1; -
DR FlyBase; FB900037131; CG14564.
SQ SEQUENCE 274 AA; 29378 MW; 9F8D2CA04755EDB8 CRC64;

Query Match 50.0%; Score 9; DB 5; Length 274;
Best Local Similarity 18.2%; Pred. No. 9,7e-14;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
XX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Koehler P., Hempel S., Entian K.-D., Hohenseil J., Jesse T.,
RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Beyan M.;
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL031135; CA20046.1; -
DR EMBL; AL161588; CAB81485.1; -
KW Hypothetical protein.
SQ SEQUENCE 303 AA; 32828 MW; AF301B0FE04AF13C CRC64;

Query Match 50.0%; Score 9; DB 10; Length 303;
Best Local Similarity 18.2%; Pred. No. 1e-13;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE Atg37380 protein.
GN Atg37380.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
XX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA.
RX MEDLINE-20084487; PubMed-10417197;
RL Lin X., Kaul S., Kounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Hearnstead M.E., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.-J., Honning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Varshen S.E., Umayam L., Tallon L.J., Gill J.R.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Ventler J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA.
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC005896; AAC98057.1; -
SQ SEQUENCE 321 AA; 34973 MW; 7E76DD64A49BBAEB CRC64;

Query Match 50.0%; Score 9; DB 10; Length 321;
Best Local Similarity 18.2%; Pred. No. 1,1e-13;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

Db 101 LASSSTATE 111

RESULT 5

Q9FH69

PRELIMINARY;

PRT: 326 AA.

AC Q9FH69;

01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE Genomic DNA, chromosome 5, TAC clone:K16E1.

Arabidopsis thaliana (Mouse ear cross).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids 11; Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702;

OX NCBI_TaxID=3702;

[1]

SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA;

MEDLINE=20181125; PubMed=10718197;

RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,

Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence

features of the regions of 3,076,755 bp covered by sixty P1 and TAC

RT clones.";

DNA Res. 7:31-63(2000).

RL EMBL; AB022210; BAB09327.1;

InterPro: IPR004345; TB2_DPL_HVA22.

DR Pfam: PF03134; TB2_DPL_HVA22; 1.

SEQUENCE 326 AA; 37321 MW; 85A3C1240B73506B CRC64;

Query Match

Best Local Similarity 50.0%; Score 9; DB 10; Length 326;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11

Db 212 LSSSSSSSE 222

RESULT 6

Q9HOC7

PRELIMINARY;

PRT: 466 AA.

AC Q9HOC7;

01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE Hypothetical 51.6 kDa protein.

DKFZP434B0535.

GN Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Muteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

OX NCBI_TaxID=9606;

[1]

RP SEQUENCE FROM N.A.

TISSUE-TESTIS;

RC MEDLINE=21154917; PubMed=11230166;

Wieman S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,

RA Ansgore W., Boecker M., Bloeker H., Bauersachs S., Bjum H.,

Lauber J., Duesterhoef A., Beyer A., Koehler K., Strack N.,

RA Meves H.W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,

Wambolt R., Korn B., Klein M., Poustka A.

RT "Towards a catalog of Human Genes and Proteins: Sequencing and

Analysis of 500 Novel Complete Protein Coding Human cDNAs."

RL Genome Res. 11:422-435(2001).

EMBL; AL136851; CAB66785.1;

DR EMBL; AL136851; CAB66785.1;

Hypothetical protein

KW SEQUENCE 466 AA; 51621 MW; 8B54AD1940CBE24 CRC64;

Query Match

Best Local Similarity 50.0%; Score 9; DB 4; Length 466;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11

Db 28 LAAAAAAAAAE 38

RESULT 7

Q96T25

PRELIMINARY;

PRT: 639 AA.

AC Q96T25;

01-DEC-2001 (TREMblrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

01-JUN-2002 (TREMblrel. 21, Last annotation update)

DE Zinc family member 5 protein.

ZINC5.

GN Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

OX NCBI_TaxID=9606;

[1]

RP SEQUENCE FROM N.A.

Gou D.M., Li W.X., Gao L., Sun Y.;

RT "A novel human zinc finger gene, hZinc5."

Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF378304; AAK5418.1;

InterPro: IPR000822; Znf_C2H2.

DR Pfam; PF00096; Zf_C2H2; 4.

PRINTS; PR01582; KV3CHANNEL.

DR Prodom; PD000003; Znf_C2H2; 1.

PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.

DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.

DNA-binding; Zinc-finger.

KW SEQUENCE 639 AA; 65849 MW; FB1E96ED28D38BF CRC64;

Query Match

Best Local Similarity 50.0%; Score 9; DB 4; Length 639;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11

Db 261 LAAAAAAAAAE 271

RESULT 8

Q9BHY3

PRELIMINARY;

PRT: 761 AA.

AC Q9BHY3;

01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE Hypothetical 80.2 kDa protein.

L5213.09.

GN Leishmania major.

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OC NCBI_TaxID=5664;

[1]

RP SEQUENCE FROM N.A.

STRAIN-FRIEDLIN;

RC MEDLINE=98146435; PubMed=9477341;

Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,

RA Smith D.F.;

"A physical map of the Leishmania major Friedlin genome."

RL Genome Res. 8:135-143(1998).

EMBL; AL583932; CAC32257.1;

DR EMBL; AL583932; CAC32257.1;

Hypothetical protein.

KW SEQUENCE 761 AA; 80172 MW; 3D21A6BA5E0109D CRC64;

Query Match

Best Local Similarity 50.0%; Score 9; DB 5; Length 761;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11

Db 682 ISTTTTTTTE 692

RESULT 9

Q9VY9 PRELIMINARY; PRT; 803 AA.

AC Q9VY9; 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)

DE CG11727 protein.

OS CG11727.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champagne M., Pfeiffer B.D.,

RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.H.G.,

RA Abtill J.F., Aspayani A., An H.-J., Andrews-Pinnkott C., Baldwin D.,

RA Baller R.M., Basu A., Bekendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bruttler P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu S., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iboviam C.,

RA Jatali J., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson P.,

RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith V.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Styrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003486; AAF48044.1;

DR FLYBase: FBgn030299; CG11727.

DR InterPro: IPR000515; BPD.transp.

DR InterPro: IPR00195; RabGAP_TBC.

DR Pfam: PF00566; TBC; 1.

DR SMART: SM00164; TBC; 1.

DR PROSITE: PS00402; BPD_TRANS_P. INN_MEMBERS: UNKNOWN_1.

SQ SEQUENCE 803 AA; 92352 MW; 204F968E447D021B CRC64;

Query Match 50.0%; Score 9; DB 5; Length 803;

Best Local Similarity 18.2%; Pred. No. 1,7e-13;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11

Db 3 LTTTTPASSAA 13

RESULT 10

Q90V2 PRELIMINARY; PRT; 1133 AA.

AC Q90V2; 01-AUG-1998 (TREMblrel. 07, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE KIA00540 protein (Fragment).

GN KIA00540.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=96290545; PubMed=9628581;

RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,

RA Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

RL DNA Res. 5:31-39(1998).

DR EMBL: AB011112; BAA25466.2;

DR InterPro: IPR000409; Beige_BEACH.

DR Pfam: PF02138; Beach; 1.

DR ProDom: PD007848; Beige_BEACH; 1.

DR PROSITE: PSS0197; BEACH; 1.

FT NON-TER

SQ SEQUENCE 1133 AA; 125533 MW; 7B460C2D0211B884 CRC64;

Query Match 50.0%; Score 9; DB 4; Length 1133;

Best Local Similarity 18.2%; Pred. No. 2.1e-13;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11

Db 683 LAAAAAAMAE 693

RESULT 11

Q90V2 PRELIMINARY; PRT; 2207 AA.

AC Q90V2; 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE Possible MDS308 homolog (Fragment).

GN L8342.01.

OS Leishmania major.

OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5664;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIDLIN;

RX Masny D., Purnelle B., Goffeau A., Ivens A.C., Lawson D., Quail M.,

RA Rajandream M.A., Barrell B.G.;

RL Submitted (JMC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIDLIN;

RX MEDLINE=98146435; PubMed=9477341;

RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,

RA Smith D.F.;

RT "A physical map of the Leishmania major Friedlin genome.";

RL Genome Res. 8:135-145(1998).
 DR EMBL: AL122012; CAB58415.1; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD. 1.
 DR Pfam: PF00271; helicase_C. 1.
 DR SMART: SM00490; HELICC; 1.
 DR ATP-binding; Helicase.
 FT NON_TER 2207 2207
 SO SEQUENCE 2207 AA; 229262 MW; E461565DA0017B0C CRC64;

Query Match 50.0%; Score 9; DB 5; Length 2207;
 Best Local Similarity 18.2%; Pred. No. 2,9e-13;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
 DB 980 LSSSAATSTE 990

RESULT 12

Q9VC00 PRELIMINARY; PRT; 2768 AA.

AC Q9VC00; PRELIMINARY; PRT; 2768 AA.
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE CG13648 protein.
 GN CG13648
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Erosophillidae; Drosophila.
 NX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B. KELEY;
 RX MEDLINE: 1196006; PubMed:10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer V.G., Champe M., Peiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borovoy D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos K., Deup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J., Honck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegwan C.,
 RA Jalali M.E., Kalush F., Karpen G.H., Ke Z., Kension J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris G., Moshell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitslks R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003750; AAF56376.1; -.
 DR Flybase: FBgn0039257; CG13648.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00093; VWF; 4.
 DR SMART: SM00214; VWF; 4.
 DR PROSITE: PS01208; VWF; 1.
 SO SEQUENCE 2768 AA; 294032 MW; CA929A21774E4684 CRC64;

Query Match 50.0%; Score 9; DB 5; Length 2768;
 Best Local Similarity 18.2%; Pred. No. 3.3e-13;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
 DB 2558 LSSSTTSTTE 2568

RESULT 13

Q9UMS8 PRELIMINARY; PRT; 33 AA.

AC Q9UMS8; PRELIMINARY; PRT; 33 AA.
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE VHL protein (Fragment).
 GN VHL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Meyer A.J., Hernandez A., Enczmann J., Gerhart C., Wernet P.,
 RA Ackermann R.;
 RT "Detection of mutations in the VHL gene in renal cell carcinoma cell
 RT lines by temperature gradient gel electrophoresis."
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: A1010196; CAB50862.1; -.
 FT NON_TER 33
 FT NON_TER 33
 SO SEQUENCE 33 AA; 3405 MW; B27C514967AD0D67 CRC64;

Query Match 44.4%; Score 8; DB 4; Length 33;
 Best Local Similarity 18.2%; Pred. No. 1.4e-09;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
 DB 23 LARAASSTATE 33

RESULT 14

Q9LL43 PRELIMINARY; PRT; 56 AA.

AC Q9LL43; PRELIMINARY; PRT; 56 AA.
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE Actin type I (Fragment).
 GN Trebouxia jamesii.
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae;
 OC Microthamniales; Trebouxia.
 NX NCBI_TaxID=53267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kroken S., Taylor J.W.;
 RT "Phylogenetic species, reproductive mode, and specificity of Trebouxia
 RT forming lichens with the fungal genus Letharia."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF242472; AAF78028.1; -.

DR InterPro: IPR004000: Actin_like.
 PFam: PF00022; actin; 1.
 FT NON_TER 1
 FT NON_TER 56
 SQ SEQUENCE 56 AA; 6357 MW; 64F701537B21BD32 CRC64;

Query Match 44.4%; Score 8; DB 10; Length 56;
 Best Local Similarity 18.2%; Pred. No. 1.8e-09;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXE 11
 |
 Db 37 LATVSSSTE 47

RESULT 15

O9NMB8 PRELIMINARY; PRT; 60 AA.
 AC O9NMB8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DP 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE Hypothetical 6.5 kDa protein (Fragment).
 GN LM26.174.
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID-5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FRIEDLIN;
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AL160493; CAB97292.1; -.
 KW Hypothetical protein.
 FT NON_TER 60
 SQ SEQUENCE 60 AA; 6484 MW; 15A5F9C8EC19325C CRC64;

Query Match 44.4%; Score 8; DB 5; Length 60;
 Best Local Similarity 18.2%; Pred. No. 1.9e-09;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXE 11
 |
 Db 27 LSSSSAAASE 37

Search completed: January 15, 2003, 12:35:28
 Job time : 35.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 16, 2003, 09:03:43 ; Search time 1525.14 Seconds
(without alignments)
209.902 Million cell updates/sec

Title: US-09-823-649a-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXX 11

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters: --
-MODEL=frame_plus_p2n.model -DEV=xlh
-O/cpnr2.1/USPTO.spool/US09823649/runat.14012003.151002.29139/app.query.fasta.1.1393
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -DOOPCI=0 -LOOPEXT=0
-UNIT5-bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pico -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09823649.ecgn.1.1.3378-etunal.14012003.151002.29139 -NCPU=6 -ICPU=3
-NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl: *
1: gb_da: *
2: gb_hlg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *

29: em_vi: *
30: em_htg_hum: *
31: em_htg_in: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pin: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrl: *
38: em_sy: *
39: em_hlgo_hum: *
40: em_hlgo_mus: *
41: em_hlgo_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	50.0	224	AY022695	AY022695 Oryza sat
2	9	50.0	224	AY022699	AY022699 Oryza sat
3	9	50.0	224	AY023282	AY023282 Oryza sat
4	9	50.0	227	AY023204	AY023204 Oryza sat
5	9	50.0	227	AY023249	AY023249 Oryza sat
6	9	50.0	311	H0M0T7920B	L30260 Human STS U
7	9	50.0	352	AX071283	AX071283 Sequence
8	9	50.0	420	CNS01BRJ	AL10710 Botrytis
9	9	50.0	480	CNS01CAI	AL115282 Botrytis
10	9	50.0	560	AF218642	AF218642 Pseudorab
11	9	50.0	621	AF450818	AF450818 Drosophila
12	9	50.0	628	AF450787	AF450787 Drosophila
13	9	50.0	631	AF450791	AF450791 Drosophila
14	9	50.0	636	CNS01DLJ	AL116975 Botrytis
15	9	50.0	660	CNS01BJR	AL114319 Botrytis
16	9	50.0	661	AX364525	AX364525 Sequence
17	9	50.0	675	HS342288	AJ342288 Homo sapi
18	9	50.0	687	HS340476	AJ340476 Homo sapi
19	9	50.0	720	CNS01IAF	AL112857 Botrytis
20	9	50.0	776	AX053733	AX053733 Sequence
21	9	50.0	812	AY074918	AY074918 Litopenae
22	9	50.0	835	HS342543	AJ342543 Homo sapi
23	9	50.0	960	AK170205	AK170205 Sequence
24	9	50.0	960	182549	182549 Sequence
25	9	50.0	1037	HC008597	HC008597 Homo sapi
26	9	50.0	1058	HC000454	HC000454 Homo sapi
27	9	50.0	1082	AX053730	AX053730 Sequence
28	9	50.0	1249	AY091089	AY091089 Arabidops
29	9	50.0	1343	AY088194	AY088194 Arabidops
30	9	50.0	1364	AY085687	AY085687 Arabidops
31	9	50.0	1530	AB055371	AB055371 Macaca fa
32	9	50.0	1536	AY128903	AY128903 Arabidops
33	9	50.0	1571	HC011834	HC011834 Homo sapi
34	9	50.0	1583	AY099728	AY099728 Arabidops
35	9	50.0	1776	HS340163	U11686 Human calmo
36	9	50.0	1808	HS340163	AL833630 Homo sapi
37	9	50.0	1826	AX405788	AX405788 Sequence
38	9	50.0	2043	AX086813	AX086813 Sequence
39	9	50.0	2043	HS340163	AL136851 Homo sapi
40	9	50.0	2087	BC000724	BC000724 Homo sapi
41	9	50.0	2147	AK091011	AK091011 Homo sapi
42	9	50.0	2210	AY058651	AY058651 Drosophila
43	9	50.0	2303	AT026944	U26944 Arabidopsis
44	9	50.0	2406	AF378304	AF378304 Homo sapi
45	9	50.0	2406	AF378304	AF378304 Homo sapi

RESULT 1

ALIGNMENTS

AY022695 224 bp DNA linear PLN 07-FEB-2001
 LOCUS
 DEFINITION Oryza sativa microsatellite MRS5020 containing (CCG)X8, closest to marker S12158, genomic sequence.
 ACCESSION AY022695
 VERSION AY022695.1 GI:12705911
 KEYWORDS
 SOURCE Oryza sativa.
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1 (bases 1 to 224)
 Tao, N., Barbazuk, W.B., Liu, J., Wu, K. and Barry, G.F.
 Simple sequence repeats from Monsanto rice genomic sequences unpublished
 2 (bases 1 to 224)
 Tao, N., Barbazuk, W.B., Liu, J., Wu, K. and Barry, G.F.
 Direct Submission
 Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA
 Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see <http://www.rice-research.org> for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.

FEATURES
 source 1..224
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 repeat_region 1..224
 /note="microsatellite MRS5020"
 /rpt_type=tandem
 /rpt_unit=ccg 54 g 38 t

BASE COUNT 26 a 106 c 54 g 38 t

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 224
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 8 Gaps: 0

US-09-823-649A-1 (1-11) x AY022695 (1-224)

QY 1 Leu*****Glu 11
 |||
 Db 97 CTCACGCCGCCGCCGCCGCCGACAG 129

RESULT 2
 LOCUS AY022699 224 bp DNA linear PLN 07-FEB-2001
 DEFINITION Oryza sativa microsatellite MRS5024 containing (CCG)X8, genomic sequence.
 ACCESSION AY022699
 VERSION AY022699.1 GI:12705915
 KEYWORDS
 SOURCE Oryza sativa.
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1 (bases 1 to 224)
 Tao, N., Barbazuk, W.B., Liu, J., Wu, K. and Barry, G.F.
 Simple sequence repeats from Monsanto rice genomic sequences unpublished
 2 (bases 1 to 224)
 Tao, N., Barbazuk, W.B., Liu, J., Wu, K. and Barry, G.F.
 Direct Submission
 Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA

COMMENT Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see <http://www.rice-research.org> for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.

FEATURES
 source 1..224
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 repeat_region 1..224
 /note="microsatellite MRS5024"
 /rpt_type=tandem
 /rpt_unit=ccg 51 g 26 t

BASE COUNT 69 a 78 c 51 g 26 t

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 224
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 8 Gaps: 0

US-09-823-649A-1 (1-11) x AY022699 (1-224)

QY 1 Leu*****Glu 11
 |||
 Db 94 CTAGCAACGCCGCCGCCGCCGCCGACAG 126

RESULT 3
 LOCUS AY023282 224 bp DNA linear PLN 07-FEB-2001
 DEFINITION Oryza sativa microsatellite MRS5607 containing (GCC)X8, genomic sequence.
 ACCESSION AY023282
 VERSION AY023282.1 GI:12706498
 KEYWORDS
 SOURCE Oryza sativa.
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1 (bases 1 to 224)
 Tao, N., Barbazuk, W.B., Liu, J., Wu, K. and Barry, G.F.
 Simple sequence repeats from Monsanto rice genomic sequences unpublished
 2 (bases 1 to 224)
 Tao, N., Barbazuk, W.B., Liu, J., Wu, K. and Barry, G.F.
 Direct Submission
 Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA
 Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see <http://www.rice-research.org> for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.

FEATURES
 source 1..224
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 repeat_region 1..224
 /note="microsatellite MRS5607"
 /rpt_type=tandem
 /rpt_unit=gcc 101 g 23 t

BASE COUNT 25 a 75 c 101 g 23 t

ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 224
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9

Gel: Acrylamide 7%, Formamide 32%, Urea 34%

Alleles: 1.

Location/Qualifiers

1..311

ORGANISM "Homo sapiens"

/db_xref="taxon:9606"

primer_bind Complement(136..134)

/evidence-experimental

BASE COUNT 68 a 89 c 105 g 47 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 311

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 11 Gaps: 0

US-09-823-649A-1 (1-11) x HUMUT7920B (1-311)

QY 1 Leu*****Glu 11

Db 151 CTGCTACTGCGCGCTGCTGCGCGCTGAG 119

RESULT 7

LOCUS AX071283 352 bp DNA linear PAT 25-JAN-2001

DEFINITION Sequence 1755 from Patent WO0102568.

ACCESSION AX071283

VERSION AX071283.1 GI:12581634

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 352)

AUTHORS Williams,D.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J., Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D., Lamson,G., Drmanac,R., Crkjenjakov,R., Drmanac,S., Dickson,M., Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Grain,B.

TITLE Human genes and gene expression products

JOURNAL Patent: WO 0102568-A 1755 11-JAN-2001;

CHIRON CORPORATION (US); HVSQ, INC. (US)

FEATURES

Location/Qualifiers

1..352

BASE COUNT 89 a 106 c 107 g 50 t

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 352

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 6 Gaps: 0

US-09-823-649A-1 (1-11) x AX071283 (1-352)

QY 1 Leu*****Glu 11

Db 79 TTACGACGACGACGACGACGACGACGACGAG 111

RESULT 8

LOCUS CNS018R1 420 bp mRNA linear PLN 02-SEP-1999

DEFINITION Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.

ACCESSION AL110710

VERSION AL110710.1 GI:5824997

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE

ORGANISM Botryotinia fuckelliana

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.

REFERENCE 1 (bases 1 to 420)

AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

TITLE Direct Submission

JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France

REFERENCE 2 (bases 1 to 420)

AUTHORS

TITLE Genoscope.

Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES

Location/Qualifiers

1..420

ORGANISM "Botryotinia fuckelliana"

/strain="T4"

/db_xref="taxon:40559"

/note="Genoscope sequence ID : W30B091"

BASE COUNT 93 a 100 c 135 g 92 t

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 420

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 8 Gaps: 0

US-09-823-649A-1 (1-11) x CNS018R1 (1-420)

QY 1 Leu*****Glu 11

Db 108 CTCGCCACGACGACGACGACGACGACGACGCA 76

RESULT 9

LOCUS CNS01CA1 480 bp mRNA linear PLN 02-SEP-1999

DEFINITION Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.

ACCESSION AL115282

VERSION AL115282.1 GI:5829901

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckelliana.

ORGANISM Botryotinia fuckelliana

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.

REFERENCE 1 (bases 1 to 480)

AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

TITLE Direct Submission

JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France

REFERENCE 2 (bases 1 to 480)

AUTHORS

TITLE Genoscope.

Direct Submission

Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

PUBMED 11919289
 REFERENCE 2 (bases 1 to 628)
 AUTHORS Machado,C.A., Kilman,R.M., Markert,J.A. and Hey,J.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2001) Department of Genetics, Rutgers University,
 604 Allison Rd., Piscataway, NJ 08854-8082, USA
 FEATURES
 source 1..628
 /organism="Drosophila pseudoobscura"
 /strain="AFC3"
 /db_xref="taxon:7237"
 /chromosome="3"
 misc_feature 1..628
 /note="3002 locus"
 BASE COUNT 166 a 150 c 129 g 183 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 628
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 3 Gaps: 0
 US-09-823-649a-1 (1-11) x AF450787 (1-628)
 QY 1 Leu*****Glu 11
 Db 477 CTCGCCGACGACGACGACGACGACGACGACGACGAA 509
 RESULT 13
 AF450791 631 bp DNA linear INV 07-APR-2002
 LOCUS Drosophila pseudoobscura strain Flagstaff6 3002 locus sequence.
 ACCESSION AF450791
 VERSION AF450791.1 GI:20066636
 KEYWORDS
 SOURCE Drosophila pseudoobscura.
 ORGANISM Drosophila pseudoobscura
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 631)
 AUTHORS Machado,C.A., Kilman,R.M., Markert,J.A. and Hey,J.
 TITLE Inferring the History of Speciation from Multilocus DNA Sequence
 JOURNAL Data: The Case of Drosophila pseudoobscura and Close Relatives
 MOL. BIOL. EVOL. 19 (4), 472-488 (2002)
 MEDLINE 21918426
 PUBMED 11919289
 REFERENCE 2 (bases 1 to 631)
 AUTHORS Machado,C.A., Kilman,R.M., Markert,J.A. and Hey,J.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2001) Department of Genetics, Rutgers University,
 604 Allison Rd., Piscataway, NJ 08854-8082, USA
 FEATURES
 source 1..631
 /organism="Drosophila pseudoobscura"
 /strain="Flagstaff6"
 /db_xref="taxon:7237"
 /chromosome="3"
 misc_feature 1..631
 /note="3002 locus"
 BASE COUNT 166 a 152 c 129 g 184 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 631
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 3 Gaps: 0

US-09-823-649a-1 (1-11) x AF450791 (1-631)
 QY 1 Leu*****Glu 11
 Db 480 CTCGCCGACGACGACGACGACGACGACGACGACGAA 512
 RESULT 14
 CNS01DLJ/c 636 bp mRNA linear PLN 03-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.
 ACCESSION AL116975
 VERSION AL116975.1 GI:5832191
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana.
 ORGANISM Botryotinia fuckeliana
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 636)
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 REFERENCE 2 (bases 1 to 636)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage ;
 CP 5706 91057 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pSII
 vector.
 FEATURES
 source 1..636
 /organism="Botryotinia fuckeliana"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W55A012"
 BASE COUNT 143 a 167 c 178 g 147 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 636
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 8 Gaps: 0
 US-09-823-649a-1 (1-11) x CNS01DLJ (1-636)
 QY 1 Leu*****Glu 11
 Db 352 CTCGCCGACGACGACGACGACGACGACGACGACGAA 320
 RESULT 15
 CNS01DLJ/c 660 bp mRNA linear PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.
 ACCESSION AL114319
 VERSION AL114319.1 GI:5828938
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana.
 ORGANISM Botryotinia fuckeliana
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 Helotiales; Sclerotiniaceae; Botryotinia.
 REFERENCE 1 (bases 1 to 660)
 AUTHORS Bitton,F., Levis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.

TITLE Direct Submission
 JOURNAL Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 REFERENCE 2 (bases 1 to 660)
 AUTHORS Genoscope.
 JOURNAL Direct Submission
 Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage :
 CP 5706 91057 EVRY cedex - FRANCE (E-mail : segrete@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.
 FEATURES Location/Qualifiers
 source 1..660
 /organism="Botryotinia fuckeliana"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : WS8F091"
 BASE COUNT 153 a 179 c 180 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 660
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 8 Gaps: 0
 US-09-823-649a-1 (1-11) x CNS01BUR (1-660)
 QY 1 Leu*****Glu 11
 |||
 DB 387 CTCCTCACCAACGCTCCACCGCGCCACCGAA 355
 |||

Search completed: January 16, 2003, 11:02:10
 Job time : 1537.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame.p1us.p2n model

Run on: January 15, 2003, 11:21:44 ; Search time 130.143 Seconds
(without alignments)
190.345 Million cell updates/sec

Title: US-09-823-649a-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXXE 11

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -v -
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/q9a2_1/USFTO.spool/US09823649/runat_14012003_151001_29119/app.query.fasta.1.1393
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=trng -MINMATCH=0.1 -LOOPEXT=0
-DOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=D10sum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pio -NORM=ext -HRAPISE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.ecsn.1.1.0.0runat_14012003_151001_29119 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMCUT=120
-WARP_TIMCUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-WGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	9	50.0	60	24	ABN37254	Human spliced tran
2	9	50.0	261	21	AAO4346	Human secreted pro
3	9	50.0	280	20	AAV90423	EST clone DM293.
4	9	50.0	343	21	AACT17532	Human secreted pro
5	9	50.0	352	22	AAE65999	Novel human polyu
6	9	50.0	362	21	AAE65999	Novel human polyu
7	9	50.0	362	21	AAE65999	Novel human polyu
8	9	50.0	362	21	AAE65999	Novel human polyu
9	9	50.0	362	21	AAE65999	Novel human polyu
10	9	50.0	437	22	AAE65999	Novel human polyu
11	9	50.0	437	22	AAE65999	Novel human polyu
12	9	50.0	442	22	AAE65999	Novel human polyu
13	9	50.0	442	22	AAE65999	Novel human polyu
14	9	50.0	442	22	AAE65999	Novel human polyu
15	9	50.0	462	22	AAE65999	Novel human polyu
16	9	50.0	462	22	AAE65999	Novel human polyu
17	9	50.0	462	22	AAE65999	Novel human polyu
18	9	50.0	462	22	AAE65999	Novel human polyu
19	9	50.0	462	22	AAE65999	Novel human polyu
20	9	50.0	462	22	AAE65999	Novel human polyu
21	9	50.0	462	22	AAE65999	Novel human polyu
22	9	50.0	462	22	AAE65999	Novel human polyu
23	9	50.0	462	22	AAE65999	Novel human polyu
24	9	50.0	462	22	AAE65999	Novel human polyu
25	9	50.0	462	22	AAE65999	Novel human polyu
26	9	50.0	462	22	AAE65999	Novel human polyu
27	9	50.0	462	22	AAE65999	Novel human polyu
28	9	50.0	462	22	AAE65999	Novel human polyu
29	9	50.0	462	22	AAE65999	Novel human polyu
30	9	50.0	462	22	AAE65999	Novel human polyu
31	9	50.0	462	22	AAE65999	Novel human polyu
32	9	50.0	462	22	AAE65999	Novel human polyu
33	9	50.0	462	22	AAE65999	Novel human polyu
34	9	50.0	462	22	AAE65999	Novel human polyu
35	9	50.0	462	22	AAE65999	Novel human polyu
36	9	50.0	462	22	AAE65999	Novel human polyu
37	9	50.0	462	22	AAE65999	Novel human polyu
38	9	50.0	462	22	AAE65999	Novel human polyu
39	9	50.0	462	22	AAE65999	Novel human polyu
40	9	50.0	462	22	AAE65999	Novel human polyu
41	9	50.0	462	22	AAE65999	Novel human polyu
42	9	50.0	462	22	AAE65999	Novel human polyu
43	9	50.0	462	22	AAE65999	Novel human polyu
44	9	50.0	462	22	AAE65999	Novel human polyu
45	9	50.0	462	22	AAE65999	Novel human polyu

ALIGNMENTS

RESULT 1
ID AEN37254
ABN37254 standard; DNA; 60 BP.
15-JUL-2002 (first entry)
Human spliced transcript detection oligonucleotide seq ID NO:10002.
Human; mouse; rat; splice transcript; detection; RNA transcript;
splice variant; transcriptome; oligonucleotide library; ss.
Homo sapiens.
WO200210449-A2.
07-FEB-2002.

```

XX PF 20-JUL-2001; 2001WO-1B01903.
XX XX
XX XX 28-JUL-2000; 2000US-221607P.
XX PR 02-MAY-2001; 2001US-287724P.
XX XX
XX PA (COMP-) COMPUGEN INC.
XX XX
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX DR WPI; 2002-257383/30.
XX XX
XX PT New oligonucleotide libraries comprising oligonucleotides which
XX PT selectively hybridize to mRNAs transcribed from a transcription unit of
XX PT a genome, useful for detecting tissue-, pathology-, and
XX PT developmental-specific genes -
XX PS
XX PS Example 1; SEQ ID 10002; 47pp; English.
XX CC
XX CC The present invention describes oligonucleotide libraries for detecting
XX CC messenger RNAs that populate a (sub-)transcriptome, where the
XX CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX CC transcription units that populate a genome. The library comprises
XX CC several oligonucleotides, each capable of hybridizing selectively to a
XX CC set of messenger RNAs transcribed from a given transcription unit of
XX CC the genome, which encodes one or more messenger RNA splice variants.
XX CC The oligonucleotide libraries are useful for detecting mRNAs from a
XX CC biological sample, in expression profiling studies, in qualitatively or
XX CC quantitatively characterizing the corresponding transcriptome, and in
XX CC detecting RNA transcripts and splice variants of human or animal
XX CC transcriptomes. The libraries may also be used as specialised mini
XX CC libraries to detect transcripts of a sub-transcriptome under a
XX CC particular biological, or pathological state, and so allowing the
XX CC detection of tissue- and pathology-specific genes such as those genes
XX CC only expressed in specific tissue under a specific pathological
XX CC condition; to detect developmental specific genes; and to detect RNA
XX CC transcripts and splice variants of a transcriptome of a patient suffering
XX CC from a particular disorder. ABN27253 to ABN59589 represent
XX CC oligonucleotide sequences from rats, humans and mice, which are used in
XX CC the exemplification of the present invention.
XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ
XX SQ Sequence 60 BP; 14 A; 20 C; 21 G; 5 T; 0 other;

Alignment Scores:
Pred. No.: 9 34 Length: 60
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 24 Gaps: 0

US-09-823-649A-1 (1-11) x ABN37254 (1-60)
QY 1 Leu*****Glu 11
DB 23 TTAGCAGCAGCAGCAGCAGCGCATCGCAG 55

RESULT 2
AAC04346/c
ID AAC04346 standard; cDNA; 261 BP.
XX AC AAC04346;
XX XX
XX DT 06-OCT-2000 (first entry)
XX XX
XX DE Human secreted protein 5' EST, SEQ ID NO: 8421.
XX XX
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.

```

```

OS Homo sapiens.
XX XX
XX PN EP1033401-A2.
XX XX
XX PD 06-SEP-2000.
XX XX
XX PF 21-FEB-2000; 2000EP-0200610.
XX XX
XX PR 26-FEB-1999; 99US-0122487.
XX XX
XX PA (GESP) GENSET.
XX XX
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR WPI; 2000-500381/45.
XX XX
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS
XX PS Claim 1; SEQ ID 8421; 71pp + CD-ROM; English.
XX CC
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors.
XX SQ
XX SQ Sequence 261 BP; 55 A; 81 C; 73 G; 50 T; 2 other;

Alignment Scores:
Pred. No.: 15 Length: 261
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 21 Gaps: 0

US-09-823-649A-1 (1-11) x AAC04346 (1-261)
QY 1 Leu*****Glu 11
DB 62 TTAGCTCCACACACCTCTCTCTCTCTCTCGCAG 30

RESULT 3
AAV90423/c
ID AAV90423 standard; cDNA; 280 BP.
XX AC AAV90423;
XX XX
XX DT 15-FEB-1999 (first entry)
XX XX
XX DE EST clone DM293.
XX XX
XX KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
XX KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
XX KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
XX KW gene therapy; ss.
XX OS Homo sapiens.
XX XX
XX PN WO9845436-A2.
XX XX
XX PD 15-OCT-1998.

```

XX 10-APR-1998; 98MO-US06955.
 XX 10-APR-1997; 97US-0838821.
 XX (GENE) GENETICS INST INC.
 XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
 PI Racie LA, Spaulding V, Treacy M;
 DR WPI: 1999-070077/06.
 XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.
 XX Claim 1; Page 536; 618pp; English.
 XX The present sequence represents a human expressed sequence tag (EST).
 CC The polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC adherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 XX Sequence 280 BP; 109 A; 60 C; 65 G; 46 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 16.4 Length: 280
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 20 Gaps: 0
 US-09-823-649A-1 (1-11) x AAV90423 (1-280)
 QY 1 Leu*****Glu 11
 DB 150 TTAGCAGCCCTTCCTCTCTTCATCAGCAG 118
 RESULT 4
 AAC17532
 ID AAC17532 standard; cDNA; 343 BP.
 XX AAC17532;
 AC
 XX 06-OCT-2000 (first entry)
 DT
 XX Human secreted protein 5' EST. SEQ ID NO: 21607.
 DE
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 KM
 XX Homo sapiens.
 OS
 XX EP1033401-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 21-FEB-2000; 2000EP-0200610.
 PF
 XX 26-FEB-1999; 99US-0122487.
 PR
 XX (GENE) GENSET.
 PA
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 PI

XX WPI: 2000-500381/45.
 DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT Claim 1; SEQ ID 21607; 71pp + CD-ROM; English.
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX Sequence 343 BP; 73 A; 84 C; 79 G; 106 T; 1 other;
 SQ
 Alignment Scores:
 Pred. No.: 17.7 Length: 343
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-823-649A-1 (1-11) x AAC17532 (1-343)
 QY 1 Leu*****Glu 11
 DB 18 TTAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 50
 RESULT 5
 AAF65999
 ID AAF65999 standard; cDNA; 352 BP.
 XX AAF65999;
 AC
 XX 09-APR-2001 (first entry)
 DT
 XX Novel human polynucleotide; SEQ ID NO: 1755.
 DE
 XX Human; cytotlastic; gene therapy; colon cancer; prostate cancer;
 KW breast cancer; lung cancer; cancer detection; ss.
 KM
 XX Homo sapiens.
 OS
 XX WO200102568-A2.
 PN
 XX 11-JAN-2001.
 PD
 XX 30-JUN-2000; 2000MO-US18374.
 PF
 XX 02-JUL-1999; 99US-0142310.
 PR
 XX 02-JUL-1999; 99US-0142311.
 PA
 XX (CHIR) CHIRON CORP.
 PA
 XX (HYSE-) HYSEQ INC.
 PI
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Cirenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;
 DR WPI: 2001-091805/10.

Library of polynucleotides for diagnosing a cancerous state of a mammalian cell and detecting cancer, particularly of the colon or prostate, comprises 3351 human polynucleotide sequences -

Claim 9; Page 795; 1046pp; English.

The present sequence is one of 3351 sequences in a library of human polynucleotides. The library is used to detect differentially expressed genes correlated with a cancerous state of a mammalian cell and can detect colon, prostate, breast and lung cancer. The library can be used to produce probes for detection of mRNA and to produce additional copies of the polynucleotides. The probes can be used for chromosome mapping of the polynucleotide and for detection of transcription levels. Ribozymes or antisense oligonucleotides can be generated. The polynucleotides and their gene products are used as genetic or biochemical markers (e.g. in blood or tissues) that will detect the earliest changes along the carcinogenesis pathway and/or monitor the efficacy of therapies and preventive interventions. The polynucleotides, polypeptides and antibodies against them can be used in pharmaceutical compositions to treat the cancers and proliferative disorders such as neoplasia, dysplasia and hyperplasia.

Sequence 352 BP; 89 A; 106 C; 107 G; 50 T; 0 other;

Alignment Scores:

Pred. No.:	17.9	Length:	352
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-1 (1-11) x AAC57599 (1-352)

Qy 1 Leu*****Glu 11
|||

Db 79 TTAGCAGCAGCAGCAGCAGCAGCATCGAG 111

RESULT 6

ID AAC57589 standard; DNA; 362 BP.

AC AAC57589;

DT 25-JAN-2001 (first entry)

DE Arachidonic acid metabolism related genomic diallelic marker #223.

Human: diallelic marker; arachidonic acid metabolism; genotyping; detection; hybridisation; phenotype; haplotype; SNP; polymorphic base; single nucleotide polymorphism; hybridisation assay; sequencing assay; specific amplification assay; identification; ERM; 12-LO-RBW; eicosanoid-related diallelic marker; 12-LO-related diallelic marker; ds.

OS Homo sapiens.

PN WO200047771-A2.

PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-IB00184.

PR 12-FEB-1999; 99US-0119917.

PR 23-MAR-1999; 99US-0275267.

PR 07-MAY-1999; 99US-0133200.

PA (GEST) GENSET.

PI Blumenfeld M, Bougueleret L, Chumakov I;

XX WPI; 2000-571881/53.

Novel diallelic markers useful for detecting conditions and genotypes associated with arachidonic acid metabolism -

Claim 13; Page 422; 802pp; English.

The present invention describes polynucleotides including diallelic markers derived from genes involved in arachidonic acid metabolism and from genomic regions flanking those genes. Methods from the present invention may be used to select individuals for clinical trials and predict responses to treatment with drugs. The polynucleotides may be used in hybridisation assays, sequencing assays and specific amplification assays for identifying an eicosanoid-related diallelic marker (ERM) or 12-LO-related diallelic marker, and for amplifying a segment of nucleotides containing an ERM. The polynucleotides are useful in diagnostic kits. The markers may be used to detect conditions and genotypes associated with arachidonic acid metabolism. AAC57367 to AAC58018 and AAB24019 and AAB24020 represent sequences used in the exemplification of the present invention.

N.B. Polymorphic bases (single nucleotide polymorphisms also known as SNPs) in the polynucleotide sequences from the present invention have been given as their corresponding degenerate bases e.g. a polymorphic base of C or T has been given as Y.

Sequence 362 BP; 51 A; 102 C; 99 G; 101 T; 0 other;

Pred. No.:	18.1	Length:	362
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	21	Gaps:	0

US-09-823-649A-1 (1-11) x AAC57589 (1-362)

Qy 1 Leu*****Glu 11
|||

Db 56 TTAGCAGCAGCAGCAGCAGCAGCATCGAG 88

RESULT 7

ID AAC57590 standard; DNA; 362 BP.

AC AAC57590;

DT 25-JAN-2001 (first entry)

DE Arachidonic acid metabolism related genomic diallelic marker #224.

Human: diallelic marker; arachidonic acid metabolism; genotyping; detection; hybridisation; phenotype; haplotype; SNP; polymorphic base; single nucleotide polymorphism; hybridisation assay; sequencing assay; specific amplification assay; identification; ERM; 12-LO-RBW; eicosanoid-related diallelic marker; 12-LO-related diallelic marker; ds.

OS Homo sapiens.

PN WO200047771-A2.

PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-IB00184.

PR 12-FEB-1999; 99US-0119917.

PR 23-MAR-1999; 99US-0275267.

PR 07-MAY-1999; 99US-0133200.

PA (GEST) GENSET.

PI Blumenfeld M, Bougueleret L, Chumakov I;

XX WPI; 2000-571881/53.

PT Novel biallelic markers useful for detecting conditions and genotypes
 associated with arachidonic acid metabolism -
 XX
 PS Claim 13; Page 422; 802pp; English.
 XX
 CC The present invention describes polynucleotides including biallelic
 CC markers derived from genes involved in arachidonic acid metabolism and
 CC from genomic regions flanking those genes. Methods from the present
 CC invention may be used to select individuals for clinical trials and
 CC predict responses to treatment with drugs. The polynucleotides may be
 CC used in hybridisation assays, sequencing assays and specific
 CC amplification assays for identifying an eicosanoid-related biallelic
 CC marker (ERBM) or 12-LO-related biallelic marker, and for amplifying a
 CC segment of nucleotides containing an ERBM. The polynucleotides are
 CC useful in diagnostic kits. The markers may be used to detect conditions
 CC and genotypes associated with arachidonic acid metabolism. AAC57367 to
 CC AAC58018 and AAB24019 and AAB24020 represent sequences used in the
 CC exemplification of the present invention.
 CC N.B. Polymorphic bases (single nucleotide polymorphisms also known as
 CC SNPs) in the polynucleotide sequences from the present invention have
 CC been given as their corresponding degenerate bases e.g. a polymorphic
 CC base of C or T has been given as Y.
 XX
 SQ Sequence 362 BP; 50 A; 102 C; 99 G; 101 T; 10 other;
 Alignment Scores:
 Pred. No.: 18.1 Length: 362
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-823-649A-1 (1-11) x AAC57590 (1-362)
 QY 1 Leu*****Glu 11
 |||
 DB 56 TTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 88
 AAC57591
 ID AAC57591 standard; DNA; 362 BP.
 XX
 AC AAC57591;
 XX
 DT 25-JAN-2001 (first entry)
 XX
 DE Arachidonic acid metabolism related genomic biallelic marker #225.
 XX
 KW Human; biallelic marker; arachidonic acid metabolism; genotyping;
 KW detection; hybridisation; phenotype; haplotype; SNP; polymorphic bases;
 KW single nucleotide polymorphism; hybridisation assay; sequencing assay;
 KW specific amplification assay; identification; ERBM; 12-LO-RBM;
 KW eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO20047771-A2.
 XX
 PD 17-AUG-2000.
 XX
 PF 11-FEB-2000; 2000MO-IB00184.
 XX
 PR 12-FEB-1999; 99US-0119917.
 PR 23-MAR-1999; 99US-0275267.
 PR 07-MAY-1999; 99US-0133200.
 XX
 PA (GIST) GENSEP.
 XX
 PI Blumenfeld M, Bougueleret L, Chumakov I;
 XX
 DR WPI; 2000-571881/53.
 XX

PT Novel biallelic markers useful for detecting conditions and genotypes
 associated with arachidonic acid metabolism -
 XX
 PS Claim 13; Page 423; 802pp; English.
 XX
 CC The present invention describes polynucleotides including biallelic
 CC markers derived from genes involved in arachidonic acid metabolism and
 CC from genomic regions flanking those genes. Methods from the present
 CC invention may be used to select individuals for clinical trials and
 CC predict responses to treatment with drugs. The polynucleotides may be
 CC used in hybridisation assays, sequencing assays and specific
 CC amplification assays for identifying an eicosanoid-related biallelic
 CC marker (ERBM) or 12-LO-related biallelic marker, and for amplifying a
 CC segment of nucleotides containing an ERBM. The polynucleotides are
 CC useful in diagnostic kits. The markers may be used to detect conditions
 CC and genotypes associated with arachidonic acid metabolism. AAC57367 to
 CC AAC58018 and AAB24019 and AAB24020 represent sequences used in the
 CC exemplification of the present invention.
 CC N.B. Polymorphic bases (single nucleotide polymorphisms also known as
 CC SNPs) in the polynucleotide sequences from the present invention have
 CC been given as their corresponding degenerate bases e.g. a polymorphic
 CC base of C or T has been given as Y.
 XX
 SQ Sequence 362 BP; 51 A; 101 C; 99 G; 101 T; 10 other;
 Alignment Scores:
 Pred. No.: 18.1 Length: 362
 Score: 9.00 Matches: 2
 Percent Similarity: 18.18% Conservative: 0
 Best Local Similarity: 18.18% Mismatches: 9
 Query Match: 50.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-823-649A-1 (1-11) x AAC57591 (1-362)
 QY 1 Leu*****Glu 11
 |||
 DB 56 TTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 88
 AAC57591
 ID AAC57591 standard; cDNA; 436 BP.
 XX
 AC AAC57591;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #20351.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 KW Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG20360.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity
XX Claim 1; SEQ ID NO 20351; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 436 BP; 162 A; 73 C; 85 G; 57 T; 59 other;
Alignment Scores:
Pred. No.: 19.3 Length: 436
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 23 Gaps: 0
US-09-823-649a-1 (1-11) x AAS84547 (1-436)
QY 1 Leu*****Glu 11
DB 329 TTGTCATCCTCTCTCTCTCTCATCATCAGAG 297
RESULT 10
AAL17892
ID AAL17892 standard; cDNA; 437 BP.
AC AAL17892:
XX
XX 07-DEC-2001 (first entry)
DT
XX
XX Human breast cancer expressed polynucleotide 10349.
DE
XX
XX Human: breast cancer; cell marker; cytostatic; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200151628-A2.
PN
XX
XX 19-JUL-2001.
PD
XX
XX 10-JAN-2001; 2001WO-US00798.
PF
XX
XX 14-JAN-2000; 2000US-0176077.
PR
XX 14-MAR-2000; 2000US-0189167.
PR
XX 24-MAR-2000; 2000US-0192099.
PR
XX 29-MAR-2000; 2000US-0193480.
PR
XX 15-MAY-2000; 2000US-0205230.
PR
XX 09-JUN-2000; 2000US-0211315.
PR
XX 25-JUL-2000; 2000US-0220534.
PA
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX
XX WPI; 2001-451856/48.
DR
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
XX
XX Claim 1; Page 1847; 3695pp; English.
PS
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 437 BP; 117 A; 91 C; 105 G; 124 T; 0 other;
Alignment Scores:
Pred. No.: 19.4 Length: 437
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 22 Gaps: 0
US-09-823-649a-1 (1-11) x AAL17892 (1-437)
QY 1 Leu*****Glu 11
DB 68 CTGACGACGACGACCTCTTCACATCAGAGAA 100
RESULT 11
AAL17993
ID AAL17993 standard; cDNA; 437 BP.
AC AAL17993:
XX
XX 07-DEC-2001 (first entry)
DT
XX
XX Human breast cancer expressed polynucleotide 10450.
DE
XX
XX Human: breast cancer; cell marker; cytostatic; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200151628-A2.
PN
XX
XX 19-JUL-2001.
PD
XX
XX 10-JAN-2001; 2001WO-US00798.
PF
XX
XX 14-JAN-2000; 2000US-0176077.
PR
XX 14-MAR-2000; 2000US-0189167.
PR
XX 24-MAR-2000; 2000US-0192099.
PR
XX 29-MAR-2000; 2000US-0193480.
PR
XX 15-MAY-2000; 2000US-0205230.
PR
XX 09-JUN-2000; 2000US-0211315.
PR
XX 25-JUL-2000; 2000US-0220534.
PA
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
DR
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
XX
XX Claim 1; Page 1867; 3695pp; English.
PS
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.

SO Sequence 437 BP; 117 A; 92 C; 106 G; 122 T; 0 other;

Alignment Scores:

Pred. No.:	19.4	Length:	437
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-1 (1-11) x AAL17993 (1-437)

OY 1 Leu*****Glu 11

DB 68 CTGACGACGACGCTCTTCACATCAAGTGA 100

RESULT 12

AAL09678
ID AAL09678 standard; CDNA; 442 BP.

XX AAL09678;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 2135.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US00798.

XX 14-JAN-2000; 2000US-0176077.

XX 14-MAR-2000; 2000US-0189167.

XX 24-MAR-2000; 2000US-0192099.

XX 29-MAR-2000; 2000US-0193480.

XX 15-MAY-2000; 2000US-0205230.

XX 09-JUN-2000; 2000US-0211315.

XX 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer

XX Claim 1; Page 412; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides

XX (AAL07544-AAL26789) and methods of assessing whether a patient is

XX afflicted with breast cancer by examining the correlation between the

XX expression of certain markers and the cancerous state of breast cells.

XX The polynucleotides and encoded polypeptides are potential markers for

XX detecting, diagnosing, monitoring, characterizing treating and

XX potentially preventing breast cancer. The polynucleotides and encoded

XX polypeptides are also useful for isolating compounds with cytostatic

XX activity.

XX Sequence 442 BP; 116 A; 93 C; 111 G; 122 T; 0 other;

Alignment Scores:

Pred. No.:	19.4	Length:	442
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-1 (1-11) x AAL09678 (1-442)

OY 1 Leu*****Glu 11

DB 83 CTGACGACGACGCTCTTCACATCAAGTGA 115

RESULT 13

AAL10819
ID AAL10819 standard; CDNA; 452 BP.

XX AAL10819;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 3276.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US00798.

XX 14-JAN-2000; 2000US-0176077.

XX 14-MAR-2000; 2000US-0189167.

XX 24-MAR-2000; 2000US-0192099.

XX 29-MAR-2000; 2000US-0193480.

XX 15-MAY-2000; 2000US-0205230.

XX 09-JUN-2000; 2000US-0211315.

XX 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer

XX Claim 1; Page 604; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides

XX (AAL07544-AAL26789) and methods of assessing whether a patient is

XX afflicted with breast cancer by examining the correlation between the

XX expression of certain markers and the cancerous state of breast cells.

XX The polynucleotides and encoded polypeptides are potential markers for

XX detecting, diagnosing, monitoring, characterizing treating and

XX potentially preventing breast cancer. The polynucleotides and encoded

XX polypeptides are also useful for isolating compounds with cytostatic

XX activity.

XX Sequence 452 BP; 117 A; 97 C; 113 G; 125 T; 0 other;

US-09-823-649A-1 (1-11) x AAL10819 (1-452)

QY 1 Leu*****Glu 11
|||
DB 83 CTGACAGCAGCAGCTCTCCACATCAAGTCAA 115

RESULT 14

ABAS1801/c
ID ABAS1801 standard; DNA; 462 BP.

XX AC ABAS1801;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #106.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-483447/52.

PS Claim 1; SEQ ID NO 106; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 462 BP; 131 A; 124 C; 88 G; 119 T; 0 other;

Alignment Scores:

Pred. No.: 19.8 Length: 462

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 22 Gaps: 0

US-09-823-649A-1 (1-11) x ABAS1801 (1-462)

QY 1 Leu*****Glu 11

|||

DB 288 CTGACAGCAGCAGCTCTCCACATCAAGTCAA 256

RESULT 15

ABAS1629/c

ID ABAS1629 standard; DNA; 462 BP.

XX AC ABAS1629;

DT 23-JAN-2002 (first entry)

DE Probe #95 for gene expression analysis in human heart cell sample.

KW Human; gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-488899/53.

PS Claim 1; SEQ ID NO 95; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart. The

CC present sequence is one such probe. The probes may be used for

CC predicting, measuring and displaying gene expression in samples derived

CC from the human heart via microarrays. By measuring gene expression, the

CC probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 462 BP; 131 A; 124 C; 88 G; 119 T; 0 other;

Alignment Scores:

Pred. No.: 19.8 Length: 462

Score: 9.00 Matches: 2

Percent Similarity: 18.18% Conservative: 0

Best Local Similarity: 18.18% Mismatches: 9

Query Match: 50.00% Indels: 0

DB: 22 Gaps: 0

US-09-823-649A-1 (1-11) x ABAS1629 (1-462)

QY 1 Leu*****Glu 11

|||

DB 288 CTGACAGCAGCAGCTCTCCACATCAAGTCAA 256

Search completed: January 15, 2003, 12:55:20

Job time: 135.143 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame.plus.p2n model

Run on: January 15, 2003, 12:36:21 ; Search time 26 Seconds

(without alignments)
129.748 Million cell updates/sec

Title: US-09-823-649A-1

Sequence: 18
18 IXXXXXXXXXXE 11

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09823649/runat_14012003_151002_29127/app_query.fasta.1.1393
-DB=Issued Patents_NA -QPM=fastap -SUFFIX=ini -MINMATCH=0.1 -IOOPCL=0
-IOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosome2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.@CGL1.1.35@runat_14012003_151002_29127 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LRARGQUERY -NEG SCORES=0 -WAIT -LONGJAG -DEV_TIMCUT=120
-MARR_TIMCUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOT=10 -FGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backlist1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.0	47	4	US-09-641-638-877	Sequence 877, App
2	50.0	47	4	US-09-641-638-878	Sequence 878, App
3	50.0	362	4	US-09-641-638-223	Sequence 223, App
4	50.0	362	4	US-09-641-638-224	Sequence 224, App
5	50.0	362	4	US-09-641-638-225	Sequence 225, App
6	50.0	960	1	US-08-597-495B-23	Sequence 23, App
7	50.0	960	4	US-09-068-051A-23	Sequence 23, App
8	50.0	2156	2	US-08-959-011-2	Sequence 2, App
9	50.0	2565	1	US-08-597-495B-29	Sequence 29, App
10	50.0	2565	4	US-09-068-051A-31	Sequence 31, App
11	50.0	2793	4	US-09-336-536-68	Sequence 68, App
12	50.0	2803	4	US-09-068-051A-33	Sequence 33, App

13	44.4	43	1	US-08-474-633A-94	Sequence 94, App
14	44.4	43	2	US-08-737-524B-18	Sequence 18, App
15	44.4	43	4	US-08-823-771-94	Sequence 94, App
16	44.4	43	5	PCT-US93-02480-17	Sequence 17, App
17	44.4	43	5	PCT-US95-08501-3	Sequence 3, App
18	44.4	47	4	US-09-641-638-876	Sequence 876, App
19	44.4	143	1	US-08-207-481-43	Sequence 43, App
20	44.4	143	5	PCT-US95-02689-49	Sequence 49, App
21	44.4	169	4	US-08-623-428B-3	Sequence 3, App
22	44.4	194	2	US-08-596-387B-107	Sequence 107, App
23	44.4	194	4	US-09-067-615-107	Sequence 107, App
24	44.4	194	5	PCT-US95-09816A-107	Sequence 107, App
25	44.4	215	1	US-08-207-481-44	Sequence 44, App
26	44.4	215	5	PCT-US95-02689-51	Sequence 51, App
27	44.4	236	1	US-08-594-031-146	Sequence 146, App
28	44.4	362	4	US-08-641-638-222	Sequence 222, App
29	44.4	405	4	US-09-199-637A-242	Sequence 242, App
30	44.4	420	3	US-09-128-155-10	Sequence 10, App
31	44.4	420	1	US-08-207-481-32	Sequence 32, App
32	44.4	420	5	PCT-US95-02689-34	Sequence 34, App
33	44.4	501	3	US-09-128-155-6	Sequence 6, App
34	44.4	534	3	US-09-128-155-3	Sequence 3, App
35	44.4	616	4	US-09-328-111-207	Sequence 207, App
36	44.4	765	4	US-09-199-637A-240	Sequence 240, App
37	44.4	802	4	US-09-293-625-1	Sequence 1, App
38	44.4	893	1	US-08-207-481-25	Sequence 25, App
39	44.4	893	5	PCT-US95-02689-25	Sequence 25, App
40	44.4	927	4	US-08-222-939-24	Sequence 24, App
41	44.4	989	3	US-09-128-155-1	Sequence 1, App
42	44.4	1013	5	US-08-207-481-38	Sequence 38, App
43	44.4	1013	1	PCT-US95-02689-40	Sequence 40, App
44	44.4	1059	4	US-09-576-160B-9	Sequence 9, App
45	44.4	1110	1	US-08-202-186-14	Sequence 14, App

ALIGNMENTS

US-09-641-638-877
Sequence 877, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bouguetelret, Lydie
APPLICANT: Chumakov, Ilva
TITLE OF INVENTION: COHEN, ADRIK
TITLE OF INVENTION: BIOMETRIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GENSET 05101
CURRENT APPLICATION NUMBER: US/09/641.638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 877
LENGTH: 47
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/key: allele
LOCATION: 24
OTHER INFORMATION: 10-87-74 : polymorphic base A or T
US-09-641-638-877

Alignment Scores:

51.5

Length:

47

```
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-641-638-877 (1-47)
QY 1 Leu*****Glu 11
   |||
Db 7 TTAGCAGCAGCAGCAGCAGCAGCATCGAG 39

RESULT 2
US-09-641-638-878
; Sequence 878, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET 051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIORITY FILING DATE: 2000-08-16
; PRIORITY FILING DATE: 2000-02-11
; PRIORITY FILING DATE: 2000-02-11
; PRIORITY FILING DATE: 1999-05-07
; PRIORITY FILING DATE: 1999-03-23
; PRIORITY FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 878
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 10-87-80 : polymorphic base A or G
US-09-641-638-878

Alignment Scores:
Pred. No.: 51.5 Length: 47
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-641-638-878 (1-47)
QY 1 Leu*****Glu 11
   |||
Db 1 TTAGCAGCAGCAGCAGCAGCAGCATCGAG 33

RESULT 3
US-09-641-638-223
; Sequence 223, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET 051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
```

```
; CURRENT FILING DATE: 2000-08-16
; PRIORITY FILING DATE: 2000-02-11
; PRIORITY FILING DATE: 2000-02-11
; PRIORITY FILING DATE: 1999-05-07
; PRIORITY FILING DATE: 1999-03-23
; PRIORITY FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 223
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 73
; OTHER INFORMATION: 10-87-74 : polymorphic base A or T
; NAME/KEY: misc_binding
; LOCATION: 53..72
; OTHER INFORMATION: 10-87-74.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 74..93
; OTHER INFORMATION: 10-87-74.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 1..18
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 345..362
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 61..85
; OTHER INFORMATION: 10-87-74 potential probe
; NAME/KEY: misc_feature
; LOCATION: 172..174
; OTHER INFORMATION: n-a, g, c or t
US-09-641-638-223

Alignment Scores:
Pred. No.: 100 Length: 362
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-641-638-223 (1-362)
QY 1 Leu*****Glu 11
   |||
Db 56 TTAGCAGCAGCAGCAGCAGCAGCATCGAG 88

RESULT 4
US-09-641-638-224
; Sequence 224, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET 051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIORITY FILING DATE: 2000-08-16
; PRIORITY FILING DATE: 2000-02-11
; PRIORITY FILING DATE: 2000-02-11
; PRIORITY FILING DATE: 1999-05-07
; PRIORITY FILING DATE: 1999-03-23
; PRIORITY FILING DATE: 1999-02-12
```

```

1 ORGANISM: Homo Sapiens
2 FEATURE:
3 NAME/KEY: allele
4 LOCATION: 138
5 OTHER INFORMATION: 10-87-140 : polymorphic base C or T
6 NAME/KEY: misc_binding
7 LOCATION: 118..137
8 OTHER INFORMATION: 10-87-140.misl, potential
9 NAME/KEY: misc_binding
10 LOCATION: 139..158
11 OTHER INFORMATION: 10-87-140.misz, potential complement
12 NAME/KEY: primer_bind
13 LOCATION: 1..18
14 OTHER INFORMATION: upstream amplification primer
15 NAME/KEY: primer_bind
16 LOCATION: 345..362
17 OTHER INFORMATION: downstream amplification primer, complement
18 NAME/KEY: misc_binding
19 LOCATION: 126..150
20 OTHER INFORMATION: 10-87-140 potential probe
21 NAME/KEY: misc_feature
22 LOCATION: 172..174
23 OTHER INFORMATION: n-a, g, c or t
24 US-09-641-638-225
25
26 Alignment Scores:
27 Pred. No.: 100 Length: 362
28 Score: 9.00 Matches: 2
29 Percent Similarity: 18.18% Conservative: 0
30 Best Local Similarity: 18.18% Mismatches: 9
31 Query Match: 50.00% Indels: 0
32 DB: 4 Gaps: 0
33
34 US-09-823-649A-1 (1-11) x US-09-641-638-225 (1-362)
35
36 Oy 1 Leu*****Glu 11
37 ||| |||
38 Db 56 TTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 88
39
40 RESULT 6
41 US-08-597-495B-23
42 Sequence 23, Application US/08597495B
43 Patent No. 5712369
44
45 GENERAL INFORMATION:
46 APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerd;
47 APPLICANT: Simpson, Richard J.; Nice, Edward; Moritz, R. L.;
48 APPLICANT: Calmel, B.; Ji, Hong; Burgess, Anthony W.;
49 APPLICANT: Heath, Joan K.; White, Sara J.; Johnstone, Cameron
50 TITLE OR INVENTION: Colon Cell And Colon Cancer Cell
51 NUMBER OF SEQUENCES: 29
52 CORRESPONDENCE ADDRESS:
53 ADDRESSEE: Felfe & Lynch
54 STREET: 805 Third Avenue
55 CITY: New York City
56 STATE: New York
57 COUNTRY: USA
58 ZIP: 10022
59
60 COMPUTER READABLE FORM:
61 MEDIUM TYPE: Diskette, 3.5 Inch, 360 kb storage
62 COMPUTER: IBM PS/2
63 OPERATING SYSTEM: PC-DOS
64 SOFTWARE: Wordperfect
65 CURRENT APPLICATION DATA:
66 APPLICATION NUMBER: US/08/597,495B
67 FILING DATE: 02-Feb-1996
68 CLASSIFICATION: 435
69 PRIOR APPLICATION DATA:
70 APPLICATION NUMBER: 08/511,876
71 FILING DATE: 04-Aug-1995
72 ATTORNEY/AGENT INFORMATION:
73 NAME: Hanson, No. 5712369man D.
74 REGISTRATION NUMBER: 30,946

```


REFERENCE/DOCKET NUMBER: LUD 5316.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ. ID NO.: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-597-495B-23

Alignment Scores:
Pred. No.: 137 Length: 960
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-1 (1-11) x US-08-597-495B-23 (1-960)

QY 1 Leu*****Glu 11
|||
DB 755 TTGGCAGCAGCAGCTACTGCTGCTGCTGCCAG 787

RESULT 7
US-09-068-051A-23
; Sequence 23, Application US/09068051A
; Patent No. 6291235

GENERAL INFORMATION:
APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerd;
Simpson, Richard J.; Nice, Edouard; Moritz,
Carmel, B.; Ji, Hong; Burgess, Anthony W.;
Heath, Joan K.; White, Sara J.; Johnstone, Cameron
TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
Associated Nucleic Acid Molecules, Protein And Peptides
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: MOTOPEPCT

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,051A
FILING DATE: 10-Dec-1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/597,495
FILING DATE: 02-Feb-1996
APPLICATION NUMBER: 08/511,876
FILING DATE: 04-Aug-1995

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6291235man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5316.2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958

INFORMATION FOR SEQ. ID NO.: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 23

US-09-068-051A-23

Alignment Scores:
Pred. No.: 137 Length: 960
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-068-051A-23 (1-960)

QY 1 Leu*****Glu 11
|||
DB 755 TTGGCAGCAGCAGCTACTGCTGCTGCTGCCAG 787

RESULT 8
US-08-959-011-2
; Sequence 2, Application US/08959011
; Patent No. 5932444

GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: VACUOLAR APPAR SUBUNIT AC45
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,011
FILING DATE: Herewith

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0412 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2156 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CRELNOT01
CLONE: 676592

US-08-959-011-2

Alignment Scores:
Pred. No.: 179 Length: 2156
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 2 Gaps: 0

US-09-823-649A-1 (1-11) x US-08-959-011-2 (1-2156)

QY 1 Leu*****glu 11
|||
Db 204 TTGGCGCGCGCGCGCGCGCGCGAG 236

RESULT 9
US-08-597-495B-29
; Sequence 29, Application US/08597495B
; Patent No. 5712369
; GENERAL INFORMATION:
; APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerd;
; APPLICANT: Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
; APPLICANT: Catimel, B.; Ji, Hong; Burgess, Anthony W.;
; APPLICANT: Heath, Joan K.; White, Sara J.; Johnstone, Cameron
; TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
; TITLE OF INVENTION: Associated Nucleic Acid Molecules, Protein And Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESS: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,495B
; FILING DATE: 02-Feb-1996
; CLASSIFICATION: 435, *
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/511,876
; FILING DATE: 04-Aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 371235man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5316.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2565 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-597-495B-29

Alignment Scores:
Pred. No.: 189 Length: 2565
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-1 (1-11) x US-08-597-495B-29 (1-2565)

QY 1 Leu*****glu 11
|||
Db 867 TTGGCATCATCATCTACTCTGCTGCTGCCGAG 899

RESULT 10
US-09-068-051A-31
; Sequence 31, Application US/09068051A
; Patent No. 6291235
; GENERAL INFORMATION:
; APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerd;
; APPLICANT: Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
; APPLICANT: Catimel, B.; Ji, Hong; Burgess, Anthony W.

Heath, Joan K.; White, Sara J.; Johnstone, Cameron
; TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
; TITLE OF INVENTION: Associated Nucleic Acid Molecules, Protein And Peptides
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,051A
; FILING DATE: 10-Dec-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,495
; FILING DATE: 02-Feb-1996
; APPLICATION NUMBER: 08/511,876
; FILING DATE: 04-Aug-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6291235man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5316.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 31
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2565 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 31
; US-09-068-051A-31

Alignment Scores:
Pred. No.: 189 Length: 2565
Score: 9.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-1 (1-11) x US-09-068-051A-31 (1-2565)

QY 1 Leu*****glu 11
|||
Db 867 TTGGCATCATCATCTACTCTGCTGCTGCCGAG 899

RESULT 11
US-09-336-536-68
; Sequence 68, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 68
; LENGTH: 2793
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-336-536-68

Alignment Scores:

Pred. No.:	194	Length:	2793
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	4	Gaps:	0

US-09-823-649A-1 (1-11) x US-09-336-536-68 (1-2793)

QY 1 Leu*****Glu 11

Db 1099 TTGGCAGTCATCATCTGCTGCTGCGGAG 1131

RESULT 12

US-09-068-051A-33

; Sequence 33, Application US/09068051A

; Patent No. 6291235

GENERAL INFORMATION:

APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerdi;
Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
Calmel, B.; Ji, Hong; Burgess, Anthony W.;
Heath, Joan K.; White, Sara J.; Johnstone, Cameron
TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
Associated Nucleic Acid Molecules, Protein And Peptides

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESS: Fulbright & Jaworski LLP

STREET: 666 Fifth Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/068, 051A

FILING DATE: 10-Dec-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/597,495

FILING DATE: 02-Feb-1996

APPLICATION NUMBER: 08/511,876

FILING DATE: 04-Aug-1995

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 6291235man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5316.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3168

TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 2803 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 33:

US-09-068-051A-33

Alignment Scores:

Pred. No.:	195	Length:	2803
Score:	9.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	50.00%	Indels:	0
DB:	4	Gaps:	0

US-09-823-649A-1 (1-11) x US-09-068-051A-33 (1-2803)

QY 1 Leu*****Glu 11

Db 1099 TTGGCAGTCATCATCTGCTGCTGCGGAG 1131

RESULT 13

US-08-474-633A-94

; Sequence 94, Application US/08474633A

; Patent No. 5773691

GENERAL INFORMATION:

APPLICANT: E. I. DU PONT DE NEMOURS AND

TITLE OF INVENTION: CHIMERIC GENES AND

METHODS FOR INCREASING

INCREASING THE LYSINE

AND THRONIN CONTENT

OF THE SEEDS OF PLANTS

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESS: E. I. DU PONT DE NEMOURS

AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD VERSION 2.0C

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474, 633A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: BARBARA C. SIEGEL

REGISTRATION NUMBER: 30,684

REFERENCE/DOCKET NUMBER: BB-1037-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-4931

TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 94:

SEQUENCE CHARACTERISTICS:

LENGTH: 43 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-474-633A-94

Alignment Scores:

Pred. No.:	392	Length:	43
Score:	8.00	Matches:	2
Percent Similarity:	18.18%	Conservative:	0
Best Local Similarity:	18.18%	Mismatches:	9
Query Match:	44.44%	Indels:	0
DB:	1	Gaps:	0

US-09-823-649A-1 (1-11) x US-08-474-633A-94 (1-43)

QY 1 Leu*****Glu 11

Db 1 CTGAGAGCTGCGCAGCTCAGCAGCGCGAA 33

RESULT 14

US-08-737-524B-18

; Sequence 18, Application US/08737524B

; Patent No. 5912414

GENERAL INFORMATION:

APPLICANT: CARL SAVERIO FALCO

APPLICANT: DOMINICK ANTHONY GUIDA, JR.
APPLICANT: MARY ELIZABETH HARNETT LOCKE
TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS, CHIMERIC
TITLE OF INVENTION: GENES AND METHODS FOR INCREASING
TITLE OF INVENTION: THE METHIONINE CONTENT OF THE SEEDS
TITLE OF INVENTION: OF PLANTS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,524B
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: LYNN M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1059-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
TELEX: 835420
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-737-524B-8

Alignment Scores:
Pred. No.: 392 Length: 43
Score: 8.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 44.44% Indels: 0
DB: 2 Gaps: 0

US-09-823-649A-1 (1-11) x US-08-737-524B-18 (1-43)

QY 1 Leu*****Glu 11
|||
DB 1 CTAGACGCTCGGCAAGCTCAGCAACGCGGAA 33

RESULT 15
US-08-823-771-94
Sequence 94, Application US/08823771
Patent No. 6459019
GENERAL INFORMATION:
APPLICANT: E. I. DU PONT DE NEMOURS AND
COMPANY
TITLE OF INVENTION: CHIMERIC GENES AND
METHODS FOR INCREASING
INCREASING THE LYSINE
AND THREONINE CONTENT
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.

ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD VERSION 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,771
FILING DATE: 24-Mar-1997
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/474,633
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BARBARA C. SIGGILL
REGISTRATION NUMBER: 30,684
REFERENCE/DOCKET NUMBER: BB-1037-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4931
TELEFAX: 302-773-0164
TELEX: 835420
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 94:
US-08-823-771-94

Alignment Scores:
Pred. No.: 392 Length: 43
Score: 8.00 Matches: 2
Percent Similarity: 18.18% Conservative: 0
Best Local Similarity: 18.18% Mismatches: 9
Query Match: 44.44% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-1 (1-11) x US-08-823-771-94 (1-43)

QY 1 Leu*****Glu 11
|||
DB 1 CTAGACGCTCGGCAAGCTCAGCAACGCGGAA 33

Search completed: January 15, 2003, 12:58:36
Job Time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:02:09 ; Search time 28.1429 Seconds

(without alignments)
52.083 Million cell updates/sec

Title: US-09-823-649A-2

Perfect score: 47

Sequence: 1 LSXELXIPYEE 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	44	93.6	11	23 AAB47793 Native DNA polymer
2	44	93.6	11	23 AAB47794 Native DNA polymer
3	44	93.6	11	23 AAB48259 Native DNA polymer
4	44	93.6	11	23 AAB48260 Native DNA polymer
5	44	93.6	11	23 AAB48261 Native DNA polymer
6	44	93.6	11	23 AAB48262 Native DNA polymer
7	44	93.6	11	23 AAB48263 Native DNA polymer
8	44	93.6	11	23 AAB48264 Native DNA polymer
9	44	93.6	11	23 AAB48265 Native DNA polymer
10	44	93.6	17	AAR96267 Mutant Thermus aqu

11	44	93.6	544	13 AAR23145	Mutant thermostabl
12	44	93.6	544	13 AAR23163	Mutant thermostabl
13	44	93.6	544	13 AAR23168	Mutant thermostabl
14	44	93.6	545	13 AAR23157	Mutant thermostabl
15	44	93.6	552	22 AAB09304	Thermus aquaticus
16	44	93.6	553	21 AAY44352	Thermosstable DNA d
17	44	93.6	554	16 AAR66209	Novel thermostable
18	44	93.6	559	21 AAY44353	Thermosstable DNA d
19	44	93.6	560	17 AAR99544	Thermus flavus DNA
20	44	93.6	560	17 AAR96205	DNA polymerase I e
21	44	93.6	561	18 AAM09315	Mutant Taq polymer
22	44	93.6	561	20 AAY00886	Mutant Taq polymer
23	44	93.6	561	20 AAY00888	FY2 polymerase. T
24	44	93.6	562	20 AAY00887	FY4 polymerase. T
25	44	93.6	562	20 AAY00887	Mutant Taq polymer
26	44	93.6	562	20 AAE00887	FY3 polymerase. T
27	44	93.6	562	22 AAE09305	Thermus thermophil
28	44	93.6	562	22 AAE09305	Thermus aquaticus
29	44	93.6	562	22 AAE09305	Thermus flavus DNA
30	44	93.6	597	17 AAR99543	DNA polymerase I e
31	44	93.6	604	22 AAU00574	Chimeric Caurtag DN
32	44	93.6	604	22 AAU00575	Chimeric Caurtag DN
33	44	93.6	631	13 AAR23144	Mutant thermostabl
34	44	93.6	632	13 AAR23156	Mutant thermostabl
35	44	93.6	632	13 AAR23162	Mutant thermostabl
36	44	93.6	632	13 AAR23167	Mutant thermostabl
37	44	93.6	632	23 AAE15565	Sulfolobus solfata
38	44	93.6	634	23 AAE15568	Polylysine (P) -
39	44	93.6	654	23 AAE15569	Mutant thermostabl
40	44	93.6	680	13 AAR23143	Mutant thermostabl
41	44	93.6	680	13 AAR23155	Mutant thermostabl
42	44	93.6	680	13 AAR23161	Mutant thermostabl
43	44	93.6	680	13 AAR23166	Mutant thermostabl
44	44	93.6	757	13 AAR23141	Mutant thermostabl
45	44	93.6	758	13 AAR23154	Mutant thermostabl

ALIGNMENTS

RESULT 1

ID AAB47793 standard; peptide; 11 AA.

AC AAB47793;

XX

DT 25-MAR-2002 (first entry)

XX

DE Native DNA polymerase motif #3.

XX

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;

KW Transverse transcription reaction; fluorescein; cyanine; thermocactive;

KW dye; amplification.

XX

OS Thermus sp.

XX

PN EPI152062-A2.

XX

PD 07-NOV-2001.

XX

PF 12-APR-2001; 2001BP-0109341.

XX

PR 18-APR-2000; 2000US-198336P.

XX

(HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Smith ES, Elstrom CM, Gelfand DH, Higuchi RG, Myers TW;

PI Schoenbrunner NJ, Wang AM;

XX

DR WPI, 2002-076891/11.

XX

PT Reverse transcribing an RNA, comprises performing a reverse

PT transcriptase polymerase chain reaction amplification of a mixture

ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US/07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudopleuronectes americanus
US-08-180-524-1

Query Match 38.9%; Score 7; DB 2; Length 37;
Best Local Similarity 18.2%; Pred. No. 4.5e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXE 11
Db 12 LTAANKAAAE 22

RESULT 15
US-08-180-524-8
Sequence 8, Application US/08180524
Patent No. 584537
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Hilde, Nick
APPLICANT: Kott, Edward
APPLICANT: Chicoye, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
TITLE OF INVENTION: IN YEAST
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thad Kryshak, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee

STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MSWORD Version 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180,524
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/917,216
FILING DATE:
APPLICATION NUMBER: US/07/486,333
FILING DATE: 28-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/409,217
FILING DATE: 19-SEP-1989
ATTORNEY/AGENT INFORMATION:
NAME: Kryshak, Thad
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 66-005-9234-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5707
TELEFAX: (414) 277-5591
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-180-524-8

Query Match 38.9%; Score 7; DB 2; Length 37;
Best Local Similarity 18.2%; Pred. No. 4.5e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXE 11
Db 12 LTAANKAAAE 22

Search completed: January 15, 2003, 12:38:20
Job time: 11.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:20:34 ; Search time 10.8571 Seconds

(without alignments)
20.140 Million cell updates/sec

Title: US-09-823-649A-1

Perfect score: 18

Sequence: 1 LXXXXXXXXXXE 11

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 segs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PTCT_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	44.4	136	9	US-10-095-407-11 Sequence 11, Appl
2	8	44.4	167	9	US-10-095-407-7 Sequence 7, Appl
3	8	44.4	176	9	US-09-965-528-25 Sequence 25, Appl
4	8	44.4	178	9	US-10-095-407-2 Sequence 2, Appl
5	8	44.4	178	9	US-10-139-833-11 Sequence 11, Appl
6	8	44.4	192	9	US-10-139-833-13 Sequence 13, Appl
7	8	44.4	193	9	US-10-063-547-142 Sequence 142, App
8	8	44.4	193	12	US-10-006-867-142 Sequence 6, Appl
9	8	44.4	198	10	US-09-788-963-6 Sequence 6397, App
10	8	44.4	206	9	US-09-738-626-6397 Sequence 12, Appl
11	8	44.4	218	9	US-10-139-833-12 Sequence 2, Appl
12	8	44.4	218	10	US-09-788-963-2 Sequence 12, Appl
13	8	44.4	398	10	US-09-925-301-1264 Sequence 1264, App
14	8	44.4	917	9	US-10-047-542-87 Sequence 87, Appl
15	8	38.9	62	10	US-09-864-761-35469 Sequence 35469, A
16	7	38.9	77	10	US-09-858-718-6 Sequence 6, Appl
17	7	38.9	94	10	US-09-764-877-1599 Sequence 1599, App
18	7	38.9	119	9	US-09-738-626-4316 Sequence 4316, App
19	7	38.9	147	9	US-09-738-626-6186 Sequence 6186, App

ALIGNMENTS

20	7	38.9	175	9	US-09-989-919-84	Sequence 84, Appl
21	7	38.9	191	9	US-09-738-626-3880	Sequence 3880, App
22	7	38.9	194	10	US-09-811-284-140	Sequence 140, App
23	7	38.9	196	10	US-09-858-664A-16	Sequence 16, Appl
24	7	38.9	300	9	US-09-938-269-1	Sequence 1, Appl
25	7	38.9	345	10	US-09-858-718-4	Sequence 4, Appl
26	7	38.9	356	9	US-09-738-626-6416	Sequence 6416, App
27	7	38.9	391	10	US-09-866-562-62	Sequence 62, Appl
28	7	38.9	430	10	US-09-815-242-11250	Sequence 11250, A
29	7	38.9	441	12	US-10-115-406-4	Sequence 4, Appl
30	7	38.9	475	10	US-09-826-752-14	Sequence 14, Appl
31	7	38.9	483	9	US-09-738-626-5097	Sequence 5097, App
32	7	38.9	486	9	US-10-108-605-51	Sequence 51, Appl
33	7	38.9	566	12	US-10-071-751-30	Sequence 30, Appl
34	7	38.9	610	9	US-09-738-626-6657	Sequence 6657, App
35	7	38.9	610	9	US-10-138-713-2	Sequence 2, Appl
36	7	38.9	610	10	US-09-852-118-2	Sequence 2, Appl
37	7	38.9	628	10	US-09-881-752A-144	Sequence 144, App
38	7	38.9	672	10	US-09-858-754-2	Sequence 2, Appl
39	7	38.9	834	10	US-09-826-752-4	Sequence 4, Appl
40	7	38.9	892	9	US-09-738-626-5307	Sequence 5307, App
41	7	38.9	911	10	US-09-745-763-140	Sequence 140, App
42	7	38.9	1034	10	US-09-858-754-5	Sequence 5, Appl
43	7	38.9	1302	12	US-10-000-864-2	Sequence 2, Appl
44	7	38.9	1400	10	US-09-764-176-7	Sequence 7, Appl
45	7	38.9	1493	10	US-09-858-754-3	Sequence 3, Appl

RESULT 1
US-10-095-407-11
; Sequence 11, Application US/10095407
; Patent No. US20020164330A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT FILING DATE: 2002-03-11
; PRIOR FILING DATE: 1998-07-02
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-095-407-11

Query Match 44.4%; Score 8; DB 9; Length 136;
Best Local Similarity 18.2%; Pred. No. 45;
Matches 2; Conservative 0; Mismatches 9; Indels 0;
Gaps 0;

QY 1 LXXXXXXXXXXE 11
DB 12 LASSLSASAE 22

RESULT 2
US-10-095-407-7
; Sequence 7, Application US/10095407
; Patent No. US20020164330A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-095-407-7

;; CURRENT FILING DATE: 2002-03-11
;; PRIOR APPLICATION NUMBER: US 60/091,650
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: US 60/054,646
;; PRIOR FILING DATE: 1997-08-04
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 7
;; LENGTH: 167
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-095-407-7

Query Match 44.4%; Score 8; DB 9; Length 167;
Best Local Similarity 18.2%; Pred. No. 51;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 43 LASSLSASAE 53

RESULT 3

US-09-965-528-25
;; Sequence 25, Application US/09965528
;; Publication No. US20020187523A1
;; GENERAL INFORMATION:
;; APPLICANT: INCYTE GENOMICS, INC.
;; APPLICANT: TANG, Y. Tom
;; APPLICANT: YUE, Henry
;; APPLICANT: LAL, Preeti
;; APPLICANT: BURROD, Neil
;; APPLICANT: BANDMAN, Olga
;; APPLICANT: BAUGHN, Mariah R.
;; APPLICANT: AZIMZAI, Yalda
;; APPLICANT: LY, Dzung Aina M.
;; APPLICANT: PATTERSON, Chandra
;; TITLE OF INVENTION: EXTRACELLULAR SIGNALING MOLECULES
;; FILE REFERENCE: PF-0701 USA
;; CURRENT APPLICATION NUMBER: US/09/965,528
;; PRIOR FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 60/134,949
;; PRIOR FILING DATE: 1999-05-19
;; PRIOR APPLICATION NUMBER: 60/144,270
;; PRIOR FILING DATE: 1999-07-15
;; PRIOR APPLICATION NUMBER: 60/146,700
;; PRIOR FILING DATE: 1999-07-30
;; PRIOR APPLICATION NUMBER: 60/157,508
;; PRIOR FILING DATE: 1999-10-04
;; NUMBER OF SEQ ID NOS: 55
;; SOFTWARE: PERL Program
;; SEQ ID NO 25
;; LENGTH: 176
;; TYPE: PRF
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; OTHER INFORMATION: Incyte ID No. US20020187523A1 4365383CD1
US-09-965-528-25

Query Match 44.4%; Score 8; DB 9; Length 176;
Best Local Similarity 18.2%; Pred. No. 51;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 52 LASSLSASAE 62

RESULT 4

US-10-095-407-2
;; Sequence 2, Application US/10095407
;; Patent No. US20020164330A1

;; GENERAL INFORMATION:
;; APPLICANT: Pan, Yang
;; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
;; FILE REFERENCE: 09404/052001
;; CURRENT APPLICATION NUMBER: US/10/095,407
;; PRIOR FILING DATE: 2002-03-11
;; PRIOR APPLICATION NUMBER: US 60/091,650
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: US 60/054,646
;; PRIOR FILING DATE: 1997-08-04
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 178
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-095-407-2

Query Match 44.4%; Score 8; DB 9; Length 178;
Best Local Similarity 18.2%; Pred. No. 51;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 54 LASSLSASAE 64

RESULT 5

US-10-139-833-11
;; Sequence 11, Application US/10139833
;; Publication No. US20030004106A1
;; GENERAL INFORMATION:
;; APPLICANT: Satis, Christiaan M.
;; APPLICANT: Giles, Jennifer
;; APPLICANT: Mu, Sharon X.
;; APPLICANT: Xia, Min
;; APPLICANT: Bass, Michael B.
;; APPLICANT: Craveiro, Roger
;; TITLE OF INVENTION: Interleukin-1 Receptor Antagonist-Related Molecules and
;; FILE REFERENCE: 00-1213-E
;; CURRENT APPLICATION NUMBER: US/10/139,833
;; CURRENT FILING DATE: 2002-05-06
;; PRIOR APPLICATION NUMBER: 60/170,191
;; PRIOR FILING DATE: 1999-12-10
;; PRIOR APPLICATION NUMBER: 60/188,053
;; PRIOR FILING DATE: 2000-03-09
;; PRIOR APPLICATION NUMBER: 60/194,521
;; PRIOR FILING DATE: 2000-04-04
;; PRIOR APPLICATION NUMBER: 60/195,910
;; PRIOR FILING DATE: 2000-04-10
;; PRIOR APPLICATION NUMBER: 09/724,583
;; PRIOR FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 37
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 11
;; LENGTH: 178
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-139-833-11

Query Match 44.4%; Score 8; DB 9; Length 178;
Best Local Similarity 18.2%; Pred. No. 51;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 54 LASSLSASAE 64

RESULT 6

US-10-139-833-13


```
Sequence 13, Application US/10139833
Publication No. US20030004106A1
GENERAL INFORMATION:
APPLICANT: Saris, Christiaan M.
APPLICANT: Giles, Jennifer
APPLICANT: Mu, Sharon X.
APPLICANT: Xia, Min
APPLICANT: Bass, Michael B.
APPLICANT: Craveiro, Roger
TITLE OF INVENTION: Interleukin-1 Receptor Antagonist-Related Molecules and
FILE OF INVENTION: Uses Thereof
FILE REFERENCE: 00-1213-E
CURRENT APPLICATION NUMBER: US/10/139,833
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: 60/170,191
PRIOR FILING DATE: 1999-12-10
PRIOR APPLICATION NUMBER: 60/188,053
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 60/194,521
PRIOR FILING DATE: 2000-04-04
PRIOR APPLICATION NUMBER: 60/195,910
PRIOR FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: 09/724,583
PRIOR FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 192
TYPE: PRT
ORGANISM: Homo sapiens
US-10-139-833-13

Query Match .44%; Score 8; DB 9; Length 192;
Best Local Similarity 18.2%; Pred. No. 53;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXE 11
DB 68 LASSISSAE 78

RESULT 7
US-10-063-547-142
Sequence 142, Application US/10063547
Publication No. US20020182638A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,547
CURRENT FILING DATE: 2002-05-02
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 142
LENGTH: 193
TYPE: PRT
ORGANISM: Homo sapiens
US-10-063-547-142

Query Match 44.4%; Score 8; DB 9; Length 193;
Best Local Similarity 18.2%; Pred. No. 53;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXE 11
```

```
DB 69 LASSISSAE 79

RESULT 8
US-10-006-867-142
Sequence 142, Application US/10006867
Patent No. US2002019130A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/006,867
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 60/063435
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/064215
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/086734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/086740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/086811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/086824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/086825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/086863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090688
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091628
PRIOR FILING DATE: 1998-07-02
```

;; PRIOR APPLICATION NUMBER: 60/096012
;; PRIOR FILING DATE: 1998-08-10
;; PRIOR APPLICATION NUMBER: 60/096757
;; PRIOR FILING DATE: 1998-08-17
;; PRIOR APPLICATION NUMBER: 60/096949
;; PRIOR FILING DATE: 1998-08-18
;; PRIOR APPLICATION NUMBER: 60/096959
;; PRIOR FILING DATE: 1998-08-18
;; PRIOR APPLICATION NUMBER: 60/097954
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/097971
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/097979
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/098749
;; PRIOR FILING DATE: 1998-09-01
;; PRIOR APPLICATION NUMBER: 60/099741
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099763
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099792
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099815
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100662
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100683
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100684
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100930
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101279
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: 60/101475
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101743
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/102570
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/103449
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103678
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103679
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103711
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105881
;; PRIOR FILING DATE: 1998-10-27
;; PRIOR APPLICATION NUMBER: 60/106030
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106856
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108807
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/112419
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/112422

;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/112853
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/113011
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/112854
;; PRIOR FILING DATE: 1998-12-16
;; PRIOR APPLICATION NUMBER: 60/113300
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113408
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113430
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/114223
;; PRIOR FILING DATE: 1998-12-30
;; PRIOR APPLICATION NUMBER: 60/115614
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/116527
;; PRIOR FILING DATE: 1999-01-20
;; PRIOR APPLICATION NUMBER: 60/116843
;; PRIOR FILING DATE: 1999-01-22
;; PRIOR APPLICATION NUMBER: 60/119285
;; PRIOR FILING DATE: 1999-02-09
;; PRIOR APPLICATION NUMBER: 60/119287
;; PRIOR FILING DATE: 1999-02-09
;; PRIOR APPLICATION NUMBER: 60/119525
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/120014
;; PRIOR FILING DATE: 1999-02-11
;; PRIOR APPLICATION NUMBER: 60/129122
;; PRIOR FILING DATE: 1999-04-13
;; PRIOR APPLICATION NUMBER: 60/129674
;; PRIOR FILING DATE: 1999-04-16
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/138387
;; PRIOR FILING DATE: 1999-06-09
;; PRIOR APPLICATION NUMBER: 60/144791
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/175481
;; PRIOR FILING DATE: 2000-01-11
;; PRIOR APPLICATION NUMBER: 60/191007
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/199397
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: 09/380139
;; PRIOR FILING DATE: 1998-08-25
;; PRIOR APPLICATION NUMBER: 09/311832
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 09/380137
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380138
;; PRIOR FILING DATE: 1999-08-25
;; PRIOR APPLICATION NUMBER: 09/380142

Query Match 44.4%; Score 8; DB 12; Length 193;
Best Local Similarity 18.2%; Pred. No. 53;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Caps 0;

Oy 1 LXXXXXXXE 11
Db 69 LASSISASAE 79

RESULT 9
us-09-788-963-6
; Sequence 6, Application US/09788963

```
; Patent No. US20020052473A1
; GENERAL INFORMATION:
; APPLICANT: YOUNG, PETER R.
; APPLICANT: MCDONNELL, PETER C.
; APPLICANT: KOMAR, SANJAY
; TITLE OF INVENTION: INTERLEUKIN-1 HOMOLOGUE, MAT IL-1H4
; FILE REFERENCE: GP-70607-1C1
; CURRENT APPLICATION NUMBER: US/09/788,963
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/293,625
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/452,140
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 198
; TYPE: PRN
; ORGANISM: HOMO SAPIENS
US-09-788-963-6
```

```
Query Match
Best Local Similarity 44.4%; Score 8; DB 10; Length 198;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
OY 1 LXXXXXXXXXE 11
DB 74 LASSLSASAE 84
```

```
RESULT 10
US-09-738-626-6397
; Sequence 6397, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OGHTAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: OKADA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7053
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6397
; LENGTH: 206
; TYPE: PRN
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6397
```

```
Query Match
Best Local Similarity 44.4%; Score 8; DB 9; Length 206;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
OY 1 LXXXXXXXXXE 11
DB 83 LATSATTAEE 93
```

RESULT 11

```
US-10-139-833-12
; Sequence 12, Application US/10139833
; Publication No. US20030004106A1
; GENERAL INFORMATION:
; APPLICANT: Saris, Christiaan M.
; APPLICANT: Giles, Jennifer
; APPLICANT: Ng, Sharon X.
; APPLICANT: Xia, Min
; APPLICANT: Bass, Michael B.
; APPLICANT: Cavaleiro, Roger
; TITLE OF INVENTION: Interleukin-1 Receptor Antagonist-Related Molecules and
; FILE REFERENCE: 00-1213-E
; CURRENT APPLICATION NUMBER: US/10/139,833
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/170,191
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/188,053
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: 60/194,521
; PRIOR FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: 60/195,910
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 09/724,583
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 218
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-139-833-12
```

```
Query Match
Best Local Similarity 44.4%; Score 8; DB 9; Length 218;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
OY 1 LXXXXXXXXXE 11
DB 94 LASSLSASAE 104
```

```
RESULT 12
US-09-788-963-2
; Sequence 2, Application US/09788963
; Patent No. US20020052473A1
; GENERAL INFORMATION:
; APPLICANT: YOUNG, PETER R.
; APPLICANT: MCDONNELL, PETER C.
; APPLICANT: KOMAR, SANJAY
; TITLE OF INVENTION: INTERLEUKIN-1 HOMOLOGUE, MAT IL-1H4
; FILE REFERENCE: GP-70607-1C1
; CURRENT APPLICATION NUMBER: US/09/788,963
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/293,625
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/452,140
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 218
; TYPE: PRN
; ORGANISM: HOMO SAPIENS
US-09-788-963-2
```

```
Query Match
Best Local Similarity 44.4%; Score 8; DB 10; Length 218;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
OY 1 LXXXXXXXXXE 11
DB 94 LASSLSASAE 104
```

RESULT 13
US-09-925-301-1264
Sequence 1264, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1264
LENGTH: 398
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SITE
LOCATION: (28)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1264

Query Match
Best Local Similarity 44.48; Score 8; DB 10; Length 398;
18.2%; Pred. No. 73;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11
DB 130 LSSSSSSSE 140

RESULT 14
US-10-047-542-87
Sequence 87, Application US/10047542
Patent No. US20020168367A1
GENERAL INFORMATION:
APPLICANT: LARICK, JAMES W.
APPLICANT: WYCOFF, KEITH L.
TITLE OF INVENTION: NOVEL IMMUNODHESINS FOR TREATING AND PREVENTING VIRAL
TITLE OF INVENTION: AND BACTERIAL DISEASES
FILE REFERENCE: 030905, 0004, CTPI
CURRENT APPLICATION NUMBER: US/10/047,542
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: PCT/US01/13932
PRIOR FILING DATE: 2001-04-28
PRIOR APPLICATION NUMBER: 60/200,298
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 101
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 87
LENGTH: 917
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-047-542-87

Query Match
Best Local Similarity 44.48; Score 8; DB 9; Length 917;
18.2%; Pred. No. 1,1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11
DB 282 LVATATATATASE 292

RESULT 15
US-09-864-761-35469
Sequence 35469, Application US/09864761

Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME DERIVED SINGLE EXON NUCLEIC ACID PHOSPHO USERFUL FO
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecolica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 35469
LENGTH: 62
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
OTHER INFORMATION: MAP TO A1109806.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.98
US-09-864-761-35469

Query Match
Best Local Similarity 38.98; Score 7; DB 10; Length 62;
18.2%; Pred. No. 3e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11
DB 40 LRAAQTSSAE 50

Tue Jan 21 10:06:36 2003

us-09-823-649a-1.rapb

Page 7

Search completed: January 15, 2003, 12:39:50
Job time : 11.8571 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:15:24 ; Search time 12.1429 Seconds
(without alignments)
87.086 Million cell updates/sec

Title: US-09-823-649a-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXX 11

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR:*
3: PIR:*
4: PIR:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	50.0	303	2 T04681	hypothetical prote
2	9	50.0	321	2 A84792	hypothetical prote
3	9	50.0	463	2 T04686	hypothetical prote
4	9	50.0	1113	2 T00271	hypothetical prote
5	8	44.4	140	2 JC4607	hypothetical prote
6	8	44.4	169	2 B75490	hypothetical prote
7	8	44.4	181	2 T05925	hypothetical prote
8	8	44.4	189	2 C97562	peptidyl prolyl ci
9	8	44.4	189	2 AB3783	peptidyl prolyl ci
10	8	44.4	207	2 B87370	hypothetical prote
11	8	44.4	325	2 S46760	hypothetical prote
12	8	44.4	334	2 T34326	homobox and LIM d
13	8	44.4	336	2 H72618	hypothetical prote
14	8	44.4	375	2 T23272	hypothetical prote
15	8	44.4	387	2 D75221	probable aspartate
16	8	44.4	387	2 AF3740	aspartate transami
17	8	44.4	388	2 A55597	oxyluciferin recept
18	8	44.4	397	2 F83621	cell division prot
19	8	44.4	401	2 AG3552	branched-chain ami
20	8	44.4	436	2 AD0896	conserved hypothet
21	8	44.4	439	2 E70075	transcription fact
22	8	44.4	441	1 E70075	nitrilotriacetate
23	8	44.4	458	2 A48392	alpha 2C adrenoc
24	8	44.4	458	2 A40392	alpha 2C adrenoc
25	8	44.4	458	2 A37869	alpha 2-adrenoc
26	8	44.4	458	2 A19480	alpha 2-adrenoc
27	8	44.4	472	2 T50015	transcription fact
28	8	44.4	472	2 F98356	hypothetical prote
29	8	44.4	472	2 AH2925	hypothetical prote

30	8	44.4	531	2 S41986	26S proteasome reg
31	8	44.4	691	2 D84889	probable receptor-
32	8	44.4	854	1 WMBP12	gene 12 protein -
33	8	44.4	854	1 WMBP22	gene 12 protein -
34	8	44.4	859	2 T26134	hypothetical prote
35	8	44.4	953	2 S54478	probable membrane
36	8	44.4	1203	2 F86266	probable phosphol
37	8	44.4	2231	2 S53416	SN1 protein - yea
38	8	44.4	2271	2 F90073	hypothetical prote
39	7	38.9	24	2 S39593	opacity protein op
40	7	38.9	35	2 S18226	opacity protein op
41	7	38.9	44	2 C82459	hypothetical prote
42	7	38.9	46	2 B82461	hypothetical prote
43	7	38.9	60	2 AC2981	hypothetical prote
44	7	38.9	76	2 T03200	hypothetical prote
45	7	38.9	82	1 FDPLAW	antifreeze protein

ALIGNMENTS

RESULT 1
T04681
hypothetical protein F8D20.270 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
C:Accession: T04681
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mey
submitted to the Protein Sequence Database, July 1998
A:Reference number: 215381
A:Accession: T04681
A:Molecule type: DNA
A:Residues: 1-303 <BEV>
A:Cross-references: EMBL:AL031135
A:Experimental source: cultivar Columbia; BAC clone F8D20
C:Genetics:
A:Map position: 4
A:Introns: 33/3; 129/1; 141/3; 179/2; 228/3; 274/3
A>Note: F8D20.270

Query Match 50.0%; Score 9; DB 2; Length 303;
Best Local Similarity 18.2%; Pred. No. 5.3e-11;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LXXXXXXXXXXE 11
Db 61 LSSSSSSSSSE 71

RESULT 2
A84792
hypothetical protein At2g37380 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Mar-2001
C:Accession: A84792
R:Li, X.; Kaul, S.; Bouniey, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Mofat, K.S.; Cronin, L.A.; Shen, M.; Yankau, S.E.; Umayan, L.; Talion,
euss, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402: 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84792
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <STO>
A:Cross-references: GB:AE002093; NID:94056491; PIDN:AAC98057.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g37380
A:Map position: 2

C:Superfamily: Arabidopsis thaliana hypothetical protein At2g37380

Query Match 50.0%; Score 9; DB 2; Length 321;
Best Local Similarity 18.2%; Pred. No. 5.4e-11;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 OY 1 LXXXXXXXXXXE 11
 Db 101 LASSSSSTAAAE 111

RESULT 3

T04686
 hypothetical protein F4B14.30 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
 C:Accession: T04686
 R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hohseil, J.; Meves, H.W.; Mayer, K.F.X
 submitted to the Protein Sequence Database, October 1998
 A:Reference number: Z15380
 A:Accession: T04686
 A:Molecule type: DNA
 A:Residues: 1-463 <REV>
 A:Cross-references: EMBL:AL031986
 A:Experimental source: cultivar Columbia; BAC clone F4B14
 C:Genetics:
 A:Map position: 4
 A:Introns: 33/3; 129/1; 156/1; 24-/3; 279/2
 A:Note: F4B14.30

Query Match 50.0%; Score 9; DB 2; Length 463;
 Best Local Similarity 18.2%; Pred. No. 6.4e-11;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11

Db 61 LSSSSSSSTSE 71

RESULT 4

T00271
 hypothetical protein KIA0540 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 11-Jan-2002
 C:Accession: T00271
 R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kolan, H.; Nomura, N.; Ohara, O.
 DNA Res. 5, 31-39, 1998
 A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complete
 A:Reference number: Z14086; M01D:96290545; PMID:9628581
 A:Accession: T00271
 A>Status: preliminary; translated from GB/EMBL/DDHU
 A:Molecule type: mRNA
 A:Residues: 1-1113 <NAG>
 A:Cross-references: EMBL:AF011112; NID:d1185382; PIDN:BA25466.1
 A:Experimental source: Brain
 C:Genetics:
 A:Note: KIA0540

Query Match 50.0%; Score 9; DB 2; Length 1113;
 Best Local Similarity 18.2%; Pred. No. 9.5e-11;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11

Db 683 LAAAAAAAE 693

RESULT 5

JC4607
 hydrophobin 1 precursor - Pisolithus tinctorius
 N:Alternate names: morphogenetic protein
 C:Species: Pisolithus tinctorius
 C>Date: 10-Apr-1996 #sequence_revision 24-May-1996 #text_change 18-Jul-2001
 C:Accession: JC4607; PC1145
 R:Tagu, D.; Nasse, B.; Martin, F.
 Gene 168, 93-97, 1996
 A:Title: Cloning and characterization of hydrophobins-encoding cDNAs from the ectomycor

A:Reference number: JC4607; M01D:96186962; PMID:8626073
 A:Accession: JC4607
 A:Molecule type: mRNA
 A:Residues: 1-140 <TAG>
 A:Cross-references: GB:U29605; NID:g905370; PIDN:AA049307.1; PID:g905371
 A:Accession: PC1145
 A:Molecule type: protein
 A:Residues: 115-140 <TAG>
 C:Comment: This protein is a secreted, low-molecular-weight and cysteine-rich fungal
 to host-roots, the aerial growth of the mycelium, and the ectomycorrhiza development
 C:Genetics:
 A:Gene: hydr-1
 C:Superfamily: 162 protein
 F:1-18/Domain: signal sequence #status predicted <STG>
 F:19-140/Product: hydrophobin 1 #status predicted <MAT>

Query Match 44.4%; Score 8; DB 2; Length 140;
 Best Local Similarity 18.2%; Pred. No. 2.7e-07;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11

Db 9 LAAAAAAVSAE 19

RESULT 6

B75490
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: B75490
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.R.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalcwski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; M01D:20036896; PMID:10567266
 A:Accession: B75490
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-169 <WHI>
 A:Cross-references: GB:AE001924; GB:AE000513; NID:g6458372; PIDN:AAF10255.1; PID:g645
 A:Experimental source: Strain R1
 C:Genetics:
 A:Gene: DR0672
 A:Map position: 1

Query Match 44.4%; Score 8; DB 2; Length 169;
 Best Local Similarity 18.2%; Pred. No. 3e-07;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXXXE 11

Db 11 LTAASSAAAE 21

RESULT 7

T05925
 hypothetical protein - barley
 C:Species: Hordeum vulgare (barley)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T05925
 R:Hess, W.R.; Golz, R.R.; Boerner, T.
 Plant Sci. 133, 191-201, 1998
 A:Title: Analysis of randomly selected cDNAs reveals the expression of stress- and d
 A:Reference number: Z15411
 A:Accession: T05925
 A>Status: preliminary; translated from GB/EMBL/DDHU
 A:Molecule type: mRNA
 A:Residues: 1-181 <HES>
 A:Cross-references: EMBL:AJ222779; NID:e1203989; PIDN:CA10984.1; PID:e1203990
 A:Experimental source: cv. Halsta, leaf

Query Match 44.4%; Score 8; DB 2; Length 181;
Best Local Similarity 18.2%; Pred. No. 3.1e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 50 LTAATAAAAVE 60

RESULT 8

C97562
peptidyl prolyl cis-trans isomerase (AF203881) [Imported] - Agrobacterium tumefaciens (9
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C/Accession: C97562
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouyollo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Matkeitz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; PMID:11743194
A/Accession: C97562
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-189 <KUR>
A/Cross-references: GB:AE007869; PIDN:AAK87452.1; PID:q15156770; GSPDB:GN00169
C/Genetics:
A/Map position: circular chromosome
C/Superfamily: peptidylprolyl isomerase; cyclophilin homology

Query Match 44.4%; Score 8; DB 2; Length 189;
Best Local Similarity 18.2%; Pred. No. 3.1e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 17 LAASTFASAAE 27

RESULT 9

AB2783
peptidyl prolyl cis-trans isomerase [Imported] - Agrobacterium tumefaciens (strain C58,
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C/Accession: AB2783
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; PMID:11743193
A/Accession: AB2783
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-189 <KUR>
A/Cross-references: GB:AE008688; PIDN:AAI42680.1; PID:q17740114; GSPDB:GN00186
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Map position: circular chromosome
C/Superfamily: peptidylprolyl isomerase; cyclophilin homology

Query Match 44.4%; Score 8; DB 2; Length 189;
Best Local Similarity 18.2%; Pred. No. 3.1e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 17 LAASTFASAAE 27

Query Match 44.4%; Score 8; DB 2; Length 207;
Best Local Similarity 18.2%; Pred. No. 3.3e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 90 LSTASASAAE 100

RESULT 10

B87370
hypothetical protein CC0974 [Imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: B87370
R/Nierman, W.C.; Feldbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: B87370
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-207 <STO>
A/Cross-references: GB:AE005673; NID:q13422256; PIDN:AAK22958.1; GSPDB:GN00148
C/Genetics:
A/Map position: circular chromosome
C/Superfamily: CCO974

RESULT 11

S46760
hypothetical protein YHR143w - yeast (Saccharomyces cerevisiae)
C/Species: Saccharomyces cerevisiae
C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 19-Apr-2002
C/Accession: S46760
R/Macri, C.
submitted to the EMBL Data Library, June 1994
A/Description: The sequence of S. cerevisiae cosmid 9666.
A/Reference number: S46752
A/Accession: S46760
A/Molecule type: DNA
A/Residues: 1-325 <MAC>
A/Cross-references: EMBL:010397; NID:q500647; PID:q500656; GSPDB:GN00008; MIPS:YHR143
C/Genetics:
A/Map position: circular chromosome
C/Superfamily: MIPS:YHR143w
A/Cross-references: SCD:S0001186
A/Map position: 8R

Query Match 44.4%; Score 8; DB 2; Length 325;
Best Local Similarity 18.2%; Pred. No. 4e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

RESULT 12

T34326
homeobox and LIM domain protein lim-6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Mar-2000
C/Accession: T34326
R/Latrelle, P.; Gattung, S.
submitted to the EMBL Data Library, April 1996
A/Description: The sequence of C. elegans cosmid K03B6.
A/Reference number: Z21506
A/Accession: T34326
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-334 <LAT>
A/Cross-references: EMBL:053375; PIDN:AAK69042.1; GSPDB:GN00028; CESP:K03B6.1
A/Experimental source: strain Bristol N2; clone K03B6

C:Genetics:
A:Gene: lhm-6: CESP:K03E6.1

A:Map position: X
A:Introns: 29/2; 65/3; 106/2; 149/3; 238/3
C:Superfamily: homeotic protein meo-3; homeobox homology; LIM metal-binding repeat homol
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 44.4%; Score 8; DB 2; Length 334;
Best Local Similarity 18.2%; Pred. No. 4e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11
DB 5 LLSATTSSTE 15

RESULT 13
H72618
hypothetical protein APE1409 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: H72618
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72618
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-336 <KAW>
A:Cross-references: DDBJ:AP000061; NID:95104821; PIDN:BAAB0406.1; PID:95105092
A:Experimental source: str2d1*KL
C:Genetics:
A:Gene: APE1409
C:Superfamily: Aeropyrum pernix hypothetical protein APE1409

Query Match 44.4%; Score 8; DB 2; Length 336;
Best Local Similarity 18.2%; Pred. No. 4.1e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11
DB 123 LASTTASSLTE 133

RESULT 14
T25272

hypothetical protein T25C8.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T25272
R:Gardner, A.
submitted to the EMBL Data Library, December 1996

A:Reference number: Z20008
A:Accession: T25272
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-375 <WIL>
A:Cross-references: EMBL:T83241; PIDN:CAB05817.1; GSPDB:GN00021; CESP:T25C8.2
A:Experimental source: clone T25C8
C:Genetics:
A:Gene: CESP:T25C8.2
A:Map position: 3
A:Introns: 167/3; 323/2
C:Superfamily: actin

Query Match 44.4%; Score 8; DB 2; Length 375;
Best Local Similarity 18.2%; Pred. No. 4.3e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11

DB 227 LMAAASSSLE 237

RESULT 15
D97521

probable aspartate transaminase (EC 2.6.1.1) aspb (similarity) [imported] - Agrobacte
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: D97521
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: D97521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK67125.1; PID:95156389; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_2460
A:Map position: circular chromosome
C:Superfamily: aspartate transaminase
C:Keywords: aminotransferase

Query Match 44.4%; Score 8; DB 2; Length 387;
Best Local Similarity 18.2%; Pred. No. 4.3e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11
DB 276 LAATAAFSAAE 286

Search completed: January 15, 2003, 12:37:06
Job time : 20.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 Seconds
(without alignments)
72.584 Million cell updates/sec

Title: US-09-823-649a-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXXE 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	9	50.0	470	1 VAS1_HUMAN	Q15904 homo sapien
2	8	44.4	140	1 HYP1_PLSTI	P52748 piscolitus
3	8	44.4	181	1 Y65L_HORVU	Q48609 hordeum vul
4	8	44.4	325	1 YH03_YEAST	P38844 saccharomyc
5	8	44.4	388	1 OXYR_MOUSE	P97926 mus muscula
6	8	44.4	388	1 OXYR_RAT	P70536 rattus norv
7	8	44.4	441	1 OXYR_BACSU	P54950 bacillus su
8	8	44.4	450	1 SWP1_ENCCU	Q9XZ71 encephalito
9	8	44.4	458	1 A2AC_MOUSE	Q01337 mus muscula
10	8	44.4	458	1 A2AC_RAT	P22086 rattus norv
11	8	44.4	531	1 RPN4_YEAST	Q03465 saccharomyc
12	8	44.4	854	1 V612_BPH2	P20345 bacterioph
13	8	44.4	854	1 V612_BPH2A	P07537 bacterioph
14	8	44.4	953	1 YMG3_YEAST	Q03516 saccharomyc
15	8	44.4	988	1 E4L2_MOUSE	Q70318 mus muscula
16	8	44.4	1005	1 E4L2_HUMAN	Q43491 homo sapien
17	8	44.4	1203	1 ALAB_ARATH	Q95415 arabidopsis
18	8	44.4	2004	1 MO2_HUMAN	Q92794 homo sapien
19	8	44.4	2231	1 SEN1_YEAST	Q00416 saccharomyc
20	8	44.4	2295	1 WDR9_HUMAN	Q95916 homo sapien
21	8	38.9	82	1 ANPA_PSEAM	P04002 pseudopieut
22	7	38.9	110	1 Y507_METTH	Q26607 methanobact
23	7	38.9	140	1 FKB2_MOUSE	P45878 mus muscula
24	7	38.9	146	1 SSI_STRYL	Q06871 streptomyce
25	7	38.9	170	1 OPR3_NEIMC	P10171 neisseria m
26	7	38.9	213	1 PYRE_PSEAE	P50587 pseudomonas
27	7	38.9	215	1 3F_DICDI	P13475 dictyosteli
28	7	38.9	240	1 SWS1_MOUSE	Q98490 mus muscula
29	7	38.9	244	1 OSU3_ARATH	P50700 arabidopsis
30	7	38.9	247	1 PRB4_HUMAN	P10163 homo sapien
31	7	38.9	248	1 PARH_VIBHA	P52626 vibrio harv
32	7	38.9	260	1 OPR1_NEIMC	P10170 neisseria m
33	7	38.9	261	1 AXIA_ARATH	Q38828 arabidopsis

34	7	38.9	267	1 Y617_SYNY3	Q55707 synechocyst
35	7	38.9	270	1 OMPG_NEIGO	P09888 neisseria g
36	7	38.9	275	1 UL34_HYEB	P28954 salmone herp
37	7	38.9	287	1 TSX_SALTY	P40776 salmonella
38	7	38.9	294	1 TSX_ENTAE	P40785 enterobacte
39	7	38.9	294	1 TSX_KLEPN	P40786 klebsiella
40	7	38.9	305	1 HOC1_HORVU	P17990 hordeum vul
41	7	38.9	308	1 CC17_YEAST	P25693 saccharomyc
42	7	38.9	308	1 Y04P_MYCTU	Q10766 mycobacteri
43	7	38.9	322	1 GPC_RAT	P16251 rattus norv
44	7	38.9	349	1 PHOE_KLEOX	Q01606 klebsiella
45	7	38.9	352	1 HEMZ_PROFR	P7183 proionibac

ALIGNMENTS

RESULT 1	ID	Sequence	Standard	PRT	470 AA.
AC	Q15904	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Vacuolar ATP synthase subunit S1 precursor (EC 3.6.3.14) (V-ATPase S1 subunit) (V-ATPase S1 accessory protein) (V-ATPase A45 subunit)				
DE	(XAP-3).				
GN	ATP6P1 OR ATP6S1 OR VAPPS1 OR XAP3.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96311563; PubMed=8733135;				
RA	Chen E.Y., Zolito M., Mazzarella R.A., Ciccodicola A., Chen C.-N.,				
RA	Zuo L., Heitner C., Burrough F.W., Ripetto M., Schlessinger D.,				
RA	D'Urso M.;				
RT	"Long-range sequence analysis in Xq28: thirteen known and six				
RT	candidate genes in 219.4 kb of high GC DNA between the RCP/GCP and				
RT	G6PD loci."				
RU	Hum. Mol. Genet. 5:659-668(1996).				
RN	[2]				
RP	SEQUENCE OF 97-470 FROM N.A.				
RX	TISSUE=Brain;				
RA	MEDLINE=94307726; PubMed=8034313;				
RA	Yokoi H., Hadano S., Kogi M., Kang X., Wakasa K., Ikeda J.;				
RT	"Isolation of expressed sequences encoded by the human Xq terminal				
RT	portion using microclone probes generated by laser microdissection."				
RU	Genomics 20:404-411(1994).				
RN	[3]				
RP	SEQUENCE OF 128-470 FROM N.A.				
RA	Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,				
RA	Suzuki Y., Ohtsuyoshi M., Nishi T., Shibahara T., Tanaka T.,				
RA	Nakamura Y., Isegaki T., Sugano S.;				
RT	"NEO human cDNA sequencing project."				
RT	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.				
CC	-1- FUNCTION: VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A VARIETY				
CC	OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +				
CC	H(+)(out).				
CC	-1- SUBUNIT: COMPOSED OF AT LEAST 10 SUBUNITS.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar (By				
CC	similarity).				
CC	-1- TISSUE SPECIFICITY: UBIQUITOUS.				
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				

```

CC EMBL: LA4140; NOT_ANNOTATED_CDS.
DR EMBL: D16469; BAA03938.1; ALT_INIT.
DR EMBL: AK026519; NOT_ANNOTATED_CDS.
DR Genew: HGNC:868; ATP6P1.
DR MIM: 300197; -.
KW ATP synthetase; Hydrogen ion transport; Hydrolyase; ATP-binding;
KW Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 41
FT CHAIN 42 470
FT TRANSMEM 420 440
FT CARBOHYD 170 170
FT CARBOHYD 261 261
FT CARBOHYD 273 273
FT CARBOHYD 296 296
FT CARBOHYD 303 303
FT CARBOHYD 350 350
FT CARBOHYD 357 357
FT CARBOHYD 329 229
FT CONFLICT 307 307
FT CONFLICT 335 335
SQ SEQUENCE 470 AA; 52025 MW; A71C7EF0E9D0D652 CRC64;

Query Match
Best Local Similarity 50.0%; Score 9; DB 1; Length 470;
Matches 2; Conservativity 18.2%; Pred. No. 2.2e-10;
Mismatch 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXE 11
DB 32 LAAAAAAAAAE 42

RESULT 2
HYPL_PISRT STANDARD; PRT; 140 AA.
AC P52748;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hydrophobin-1 precursor.
GN HYD-1
OS Pisolithus tinctorius.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Boletales; Sclerotiumataceae; Pisolithus.
OX NCBI_TaxID=37468;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=441;
RX MEDLINE=96186962; PubMed=8626073;
RA Tagu D., Nasse B., Martin F.;
RT "Cloning and characterization of hydrophobins-encoding cDNAs from
RT the ectomycorrhizal basidiomycete Pisolithus tinctorius.";
RL Gene 168:93-97(1996).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PUNGAL HYDROPHOBIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U29605; AAC49307.1; -.
DR InterPro: IPR001338; Hydrophobin.
DR Pfam: PF01185; Hydrophobin.1.
DR SMART: SM00075; HYDRO.1.1.
DR PROSITE: PS00956; HYDROPHOBIN.1.
KW Cell wall; Signal.
FT SIGNAL 1 18
FT CHAIN 19 140
FT SEQUENCE 140 AA; 13965 MW; 0A05E8D607DF4E33 CRC64;

```

```

Query Match
Best Local Similarity 44.4%; Score 8; DB 1; Length 140;
Matches 2; Conservativity 18.2%; Pred. No. 4.9e-07;
Mismatch 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXE 11
DB 9 LAAAAAAAAAE 19

RESULT 3
Y65L_HORVU STANDARD; PRT; 181 AA.
ID Y65L_HORVU
AC 048609;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ycf65-like protein, chloroplast precursor.
DE Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Halsea; TISSUE=leaf;
RA Hess W.R., Goltz R., Boerner T.;
RT "Analysis of randomly selected cDNAs reveals the expression of stress-
RT and defence-related genes in the barley mutant alboblasticans.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DDJ databases.
CC -1- SUBCELLULAR LOCATION: Chloroplast (Potential).
CC -1- SIMILARITY: BELONGS TO THE YCF65 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AJ222779; CA10984.1; -.
KW Hypothetical protein; Chloroplast; transit peptide.
FT TRANSIT 1 181
FT CHAIN 191
FT SEQUENCE 181 AA; 19865 MW; B02DAC3792F7E8B5 CRC64;

Query Match
Best Local Similarity 44.4%; Score 8; DB 1; Length 181;
Matches 2; Conservativity 18.2%; Pred. No. 5.6e-07;
Mismatch 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXE 11
DB 50 LAAAAAAAAAE 60

RESULT 4
YH03_YEAST STANDARD; PRT; 325 AA.
ID YH03_YEAST
AC P38844;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 33.4 kDa protein in RPL44B-RPC10 intergenic region
DE precursor.
DE YHRL43W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;

```

```

RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Galling S., Geisel C., Kirsten J.,
RA Kuwabara T., Hillier L., Jier M., Johnston L., Johnston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Man M., Rifkin L., Riles L., St Peter H., Trevisan K., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston K., Wilson K.,
RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science. 265:2077-2082(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U10397; AB68983.1; -.
CC PIR: S46760; S46760.
CC SGD: S0001186; YHR143W.
CC DR Hypothetical protein; Signal.
CC KM CHAIN 1 23 POTENTIAL.
CC FT SIGNAL 1 23 HYPOTHETICAL PROTEIN YHR143W.
CC FT CHAIN 24 325
CC SQ SEQUENCE 325 AA; 33417 MW; 72FD17D5C919A1F CRC64;
Query Match 44.4%; Score 8; DB 1; Length 325;
Best Local Similarity 18.2%; Pred. No. 7.6e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11
DB 225 LASSSTSAVE 235
RESULT 5
OXYR_MOUSE STANDARD; PRT; 388 AA.
ID OXYR_MOUSE
AC P97926;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Oxytocin receptor (OT-R).
GN OXTR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Murelocostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=97179034; PubMed=9027321;
RA Kubota Y., Kimura T., Hashimoto K., Tokugawa Y., Nobunaga K.,
RA Azuma C., Saij F., Murata Y.,
RT "Structure and expression of the mouse oxytocin receptor gene."
RL Mol. Cell. Endocrinol. 124:25-32(1996).
CC -1- FUNCTION: RECEPTOR FOR OXYTOCIN. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D86599; BAA18995.1; -.

```

```

DR EMBL: D86631; BAA18996.1; -.
DR MGD; MGI:109147; OXtr.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECPT_P1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECPT_P1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 63 1 (POTENTIAL).
FT DOMAIN 64 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 97 2 (POTENTIAL).
FT DOMAIN 98 113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 114 135 3 (POTENTIAL).
FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 4 (POTENTIAL).
FT DOMAIN 176 202 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 203 225 5 (POTENTIAL).
FT DOMAIN 226 274 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 275 293 6 (POTENTIAL).
FT DOMAIN 294 308 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 309 331 7 (POTENTIAL).
FT DOMAIN 332 388 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 8 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 112 187 BY SIMILARITY.
FT DOMAIN 129 132 POLY-LEU.
FT DOMAIN 236 242 POLY-ALA.
SQ SEQUENCE 388 AA; 42751 MW; 77FF829ACBE5285 CRC64;
Query Match 44.4%; Score 8; DB 1; Length 388;
Best Local Similarity 18.2%; Pred. No. 8.4e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11
DB 233 LKTAANAANE 243
RESULT 6
OXYR_RAT STANDARD; PRT; 388 AA.
ID OXYR_RAT
AC P70536;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxytocin receptor (OT-R).
GN OXTR OR OTR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Murelocostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95116525; PubMed=7816817;
RA Rozen F., Russo C., Banville D., Zingg H.H.;
RT "Structure, characterization, and expression of the rat oxytocin
RT receptor gene."
RL Proc. Natl. Acad. Sci. U.S.A. 92:200-204(1995).
RN [2]
RP REVISION TO 84.
RA Zingg H.H.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases
CC -1- FUNCTION: RECEPTOR FOR OXYTOCIN. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

RA Dentzov F., Devine K.M., Dustierhoff A., Ehrlich S.D., Emmerson P.T.,
RA Britton K.D., Errington J., Fiebert C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleon N.,
RA Gilm S.Y., Glaser P., Goffeau A., Goughly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henatt A.,
RA Hilbert B., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Jorts B., Karamata D., Kashara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.W., Levine A., Liu H., Masuda S., Mauel C., Médigue C.,
RA Meene D., Melliardo R.P., Mizuno M., Moesli D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potpourlik S., Prescott A.M.,
RA Prescen E., Puig C., Purteille B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scorrone F.,
RA Sekiguchi Y., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Yamakoshi A., Yanaka T., Yersita P., Yognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassartoli A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Wetzpenger T.,
RA Winters P., Wippl A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.,
RT The complete genome sequence of the Gram-positive bacterium *Bacillus*
RP subtilis".
RL Nature 390:249-256(1997).
CC
CC
CC -1- SIMILARITY: BELONGS TO THE NTAA/SNAA/SOXA(DSZA) FAMILY OF
CC MONOOXYGENASES.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: D45912; BAA08327.1; -
DR EMBL: Z99124; CAB15988.1; -
DR Subtilist; Bg11887; yxex.
KW Hypothetical protein; Oxidoreductase; Monooxygenase;
KW Complete proteome.
SQ SQUENCE 441 AA; 49339 MW; 4C7FAF8DF34C906 CRC64;

Query Match 44.4%; Score 8; DB 1; Length 441;
Best local similarity 18.2%; Pred. No. 8,9e-07;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LXXXXXXXXXXE 11
Db 93 LTRASTYSTE 103

RESULT 8
EMBL ENCCU
ID SWML ENCCU STANDARD: PRT; 450 AA.
AC Q9XZY1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Spore wall protein 1 precursor.
GN SWP1.
OS Enecephalitozon cuniculi.
OC Eukaryota; Microsporidia; Unikaryoniadae; Enecephalitozoan.
OX NCB1_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RP Bohne W., Ferguson D.J.P., Kohler K., Gross U.;
RT "Molecular characterisation of a developmentally expressed spore wall
PROT protein from the human microsporidian Enecephalitozon cuniculi".
Submitted (Mar-1993) to the EMBL/GenBank/DBD databases.
CC
CC -1- FUNCTION: SPORE WALL COMPONENT.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ33745; CAB39735.1; -
 CC
 CC Sporulation; Signal.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 450 SPORE WALL PROTEIN 1.
 FT DOMAIN 66 74 POLY-SER.
 FT DOMAIN 79 82 POLY-ARG.
 FT DOMAIN 303 339 THR-RICH.
 FT DOMAIN 355 450 GLY/SER-RICH.
 SQ SEQUENCE 450 AA: 45873 MW: 5070133E3A6DF60 CRC64;

Query Match 44.4%; Score 8; DB 1; Length 450;
 Best Local Similarity 18.2%; Pred. No. 9e-07;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
 Db 327 LTRITSTSTE 337

RESULT 9
 A2AC MOUSE STANDARD; PRT; 458 AA.
 AC 001337; Created
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Alpha-2C adrenergic receptor (Alpha-2C adrenoceptor) (Subtype C4).
 GN ADR2C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Molesteostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92342131; PubMed=1353249;
 RA Link R.E., Daut D.A., Barsh G., Chruscinski A.J., Koblika B.K.;
 RT "Cloning of two mouse genes encoding alpha 2-adrenergic receptor
 RT subtypes and identification of a single amino acid in the mouse alpha
 RT 2-C10 homolog responsible for an interspecies variation in
 RT antagonist binding.";
 RL Mol. Pharmacol. 42:16-27(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DNA/2;
 RX MEDLINE=91250567; PubMed=8387367;
 RA Chang Y.-H., Chang A.C., Chen W.-M., Chang N.-C.A.;
 RT "Molecular characterization of a murine homologue of alpha 2C4
 RT adrenoceptor subtype gene";
 RL Biochem. Mol. Biol. Int. 29:467-474(1993).
 CC -1- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
 CC INDUCED INHIBITION OF ADENYLYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M93376; AAA37212.1; -
 DR EMBL: M93716; AAA37183.1; -

DR HSDP; P29274; 1MH.
 DR MGD; MGI:87936; Adra2c.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RCDP_P1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RCDP_P2_1; 1.
 KW G-Protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Phosphorylation.
 FT DOMAIN 1 51
 FT TRANSMEM 52 76
 FT DOMAIN 77 88
 FT TRANSMEM 89 114
 FT DOMAIN 115 124
 FT TRANSMEM 125 147
 FT DOMAIN 148 168
 FT TRANSMEM 169 191
 FT DOMAIN 192 207
 FT TRANSMEM 208 231
 FT DOMAIN 232 379
 FT TRANSMEM 380 403
 FT DOMAIN 404 416
 FT TRANSMEM 417 437
 FT DOMAIN 438 458
 FT DOMAIN 491 305
 FT CARBOHYD 19 19
 FT CARBOHYD 33 33
 FT DISULFID 124 202
 FT CONFLICT 196 196
 FT CONFLICT 296 296
 FT CONFLICT 298 298
 SQ SEQUENCE 458 AA: 49906 MW: C0A8BDF032BFFB CRC64;

Query Match 44.4%; Score 8; DB 1; Length 458;
 Best Local Similarity 18.2%; Pred. No. 9e-07;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXK 11
 Db 6 LMAALAAAE 16

RESULT 10
 A2AC RAT STANDARD; PRT; 458 AA.
 AC 01-APR-1991 (Rel. 19, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Alpha-2C adrenergic receptor (Alpha 2C adrenoceptor) (Subtype C4).
 GN ADR2C.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91126047; PubMed=1704126;
 RA Floderus G.S., Handy D.E., Bresnahan M.R., Zannis V.I., Gavras H.;
 RT "Cloning and expression of a rat brain alpha 2b-adrenergic receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:1019-1023(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91130596; PubMed=1704314;
 RA Voigt M.M., McCune S.K., Kanterman R.Y., Felder C.C.;
 RT "The rat alpha 2-C4 adrenergic receptor gene encodes a novel
 RT pharmacological subtype.";
 RL PNAS Lett. 278:45-50(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91244823; PubMed=1645350;
 RA Lanier S.M., Downing S., Duzie R., Homcy C.J.;
 RT "Isolation of rat genomic clones encoding subtypes of the alpha 2

DR StyGene; SG10210; lrp.
DR InterPro: IPR000485; ASNC_trans_reg.
DR Pfam: PF01037; ASNC_trans_reg; 1.
DR SMART: SM00344; HTH_ASCNC; 1.
DR PROSITE: PS00519; HTH_ASCNC_FAMILY; 1.
KW DNA-binding; transcription regulation; Activator; Complete proteome.
FT INT_MET 0 BY SIMILARITY.
FT DNA_BIND 30 49 H-T-H MOTIF (POTENTIAL).
FT CONFLICT 85 A -> G (IN REF. 1).
SQ SEQUENCE 163 AA: 18725 MW: A01935958BED25 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 163;
Best Local Similarity 87.5%; Pred. No. 8.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLS 8
DB 33 LSKRVGLS 40

RESULT 11
LRP_SERMA STANDARD; PRT; 163 AA.
AC P37425;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Leucine-responsive regulatory protein.
GN LRP.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615; -- *
RN [1]
RP SEQUENCE FROM N.A.
RA Calvo J.M.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MEDIATES A GLOBAL RESPONSE TO LEUCINE. EXOGENOUS LEUCINE
CC AFFECTS THE EXPRESSION OF A NUMBER OF DIFFERENT OPERONS. LRP
CC MEDIATES THIS EFFECT FOR AT LEAST SOME OF THESE OPERONS. FOR
CC EXAMPLE IT IS REGULATOR OF THE BRANCHED-CHAIN AMINO ACID TRANSPORT
CC GENES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U02276; AAA75466.1; -
DR InterPro: IPR000485; ASNC_trans_reg.
DR Pfam: PF01037; ASNC_trans_reg; 1.
DR PRINTS: PR00033; HTHASCNC.
DR SMART: SM00344; HTH_ASCNC; 1.
DR PROSITE: PS00519; HTH_ASCNC_FAMILY; 1.
KW DNA-binding; transcription regulation; Activator.
FT INT_MET 0 BY SIMILARITY.
FT DNA_BIND 30 49 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 163 AA: 18725 MW: 0940AC4CT937AE48 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 163;
Best Local Similarity 87.5%; Pred. No. 8.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLS 8
DB 33 LSKRVGLS 40

RESULT 12
-GAL1_SCHPO STANDARD; PRT; 519 AA.
AC Q9HDZ2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Galactokinase (EC 2.7.1.6) (galactose kinase).
GN SPP2B2.13.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagers K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnietz B.,
RA Wellens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer C., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Weiler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode J.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Useery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- CATALYTIC ACTIVITY: ATP + D-galactose - ADP + D-galactose 1-
CC phosphate.
CC -----
CC -1- PATHWAY: Galactose metabolism; first step.
CC -1- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. GALK SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL512522; CAC21415.1; -
DR InterPro: IPR000705; Galactokinase.
DR InterPro: IPR001745; GMPkinse_ATP.
DR InterPro: IPR001459; Mew_gal_kin.
DR Pfam: PF00288; GMP_KINASES; 1.
DR PRINTS: PR00473; GALCOKINASE.
DR PRINTS: PR00959; MEYGALKINASE.
DR TIGRfam: TIGR00131; gal_kin_1.
DR PROSITE: PS00106; GALCOKINASE; 1.
DR PROSITE: PS00627; GMP_KINASES_ATP; 1.
KW Phosphatase; kinase; Galactose metabolism; ATP-binding.
FT NP_BIND 157 167 ATP (POTENTIAL).
SQ SEQUENCE 519 AA: 58385 MW: 5EB9C12986F0C40 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 519;

Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
Db 352 LANRIGLSVSE 362

RESULT 13
YMO4_SCHPO STANDARD; PRT: 585 AA.
ID YMO4_SCHPO
AC O9HG47;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C543.04 in chromosome II.
GN SPC543.04.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sounos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Saylor J., Simmonds M., Squares K., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Gymnopoulos B.,
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck C., Leinach H., Reinhardt K., Pohl T.M.,
RA Beyer P., Zimmermann W., Medler H., Mambolt R., Purnelle H.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Heullo J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Kortsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe".
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: BELONGS TO THE UPF0171 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL: AL391604; CAC05246.1; -
CC DR InterPro: IPR005365; UPF0171.
CC DR Pfam: PF03666; UPF0171; 1.
CC KM Hypothetical protein
CC SEQUENCE 585 AA; 66866 MW; 925303F55AC1B59 CR664;

Query Match 71.4%; Score 35; DB 1; Length 585;
Best Local Similarity 54.5%; Pred. No. 31;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11

Db 359 LANRIGLSVSE 369
:::1:11111

RESULT 14
RS4_BACSU STANDARD; PRT: 199 AA.
ID RS4_BACSU
AC P21466; O53282; O45662;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S4 (BS4).
GN rpsD.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=91035248; PubMed=169930;
RA Grundy F., Henkin T.M.;
RT "Cloning and analysis of the Bacillus subtilis rpsD gene, encoding
RT ribosomal protein S4.";
RL J. Bacteriol. 172:6372-6379(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rnm-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [3]
RP SEQUENCE OF 1-56 FROM N.A.
RX MEDLINE=91310564; PubMed=1906866;
RA Grundy F.D., Henkin T.M.;
RT "The rpsD gene, encoding ribosomal
RT protein S4, is autogenously
RT regulated in Bacillus subtilis.";
RL J. Bacteriol. 173:4595-4602(1991).
RN [4]
RP SEQUENCE OF 1-24.
RX MEDLINE=82219212; PubMed=6806564;
RA Higo K.I., Otake F., Osawa S.;
RT "Purification and characterization of 30S ribosomal proteins from
RT Bacillus subtilis: correlation to Escherichia coli 30S proteins.";
RL Mol. Gen. Genet. 167:239-244(1987).
RN [5]
RP MUTANTS.
RX MEDLINE=91035249; PubMed=2121712;
RA Henkin T.M., Chambliss G.H., Grundy F.;
RT "Bacillus subtilis mutants with alterations in ribosomal protein S4.";
RL J. Bacteriol. 172:6380-6385(1990).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC -1- CONTROLS THE TRANSLATION OF THE ALPHA-OPERON (WHICH CODERS FOR S13,
CC S11, S4, RNA POLYMERASE ALPHA SUBUNIT, AND IL7) BY BINDING TO ITS
CC MRNA.
CC -1- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL: M59358; AAA22717.1; -
CC DR EMBL: AF008220; AAC00397.1; -
CC DR EMBL: S45404; AAB19387.1; -
CC DR EMBL: Y99119; CAH14944.1; -
CC DR EMBL: M60889; AAA22716.1; -

```

DR PIR: A37146; A37146.
DR HSSP: P81288; 1C05.
DR Subtilist: HG10372; rpsd.
DR InterPro: IPR001912; Ribosomal_S4.
DR InterPro: IPR002942; S4.
DR Pfam: PF00163; Ribosomal_S4; 1.
DR Pfam: PF01479; S4; 1.
DR SMART: S000363; S4; 1.
DR TIGRfam: TIGR01017; rpsd_bact; 1.
DR PROSITE: PS00632; RIBOSOMAL_S4; 1.
KW Ribosomal protein; rRNA-binding; Translation regulation; Repressor;
KW RNA-binding; Complete proteome.
FT INIT_MET 0 0
FT DOMAIN 91 138 RNA-BINDING (S4 TYPE).
FT VARIANT 74 77 MISSING (IN MUTANT RPSD2).
FT VARIANT 77 77 L->IACKI (IN MUTANT RPSD1).
SQ SEQUENCE 199 AA; 22704 MW; 23D69C542E22134 CRC64;

Query Match
Best local similarity 69.4%; Score 34; DB 1; Length 199;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLSVS 10
DB 10 LSKRIGLSVS 19
|||:|:|:|:|

RESULT 15
YAS_BACSU
ID YAS_BACSU STANDARD; PRT; 201 AA.
AC P37505;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yas.
GN YAS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT *Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.*;
RL DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azeredo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortiss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleier-Bianchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Meliado R.P., Mizuno M., Mostl D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portoullik S., Prescott A.M.,
RA Prescan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Roche E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schreier R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

```

```

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tononi A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weizenecker T.,
RA Winters P., Wipet A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT *The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.*;
RL Nature 390:249-256(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D26185; BAA05204.1; -
CC EMBL: Z99124; CAB16110.1; -
CC Subtilist: HG10032; YAS.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 201 AA; 22012 MW; 77A149C6F0979317 CRC64;

Query Match
Best local similarity 77.8%; Score 34; DB 1; Length 201;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLSV 9
DB 37 LSKRIGLTV 45
|||:|:|:|:|

```

Search completed: January 15, 2003, 11:21:33
Job time : 11.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 10:29:39 ; Search time 25.5714 Seconds

(without alignments)
88.635 Million cell updates/sec

Title: US-09-823-649A-6

Perfect score: 1 LSKRIGLSVSE 11

Sequence:

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	83.7	317	17	Q58439
2	39	79.6	984	17	Q58439
3	38	77.6	352	5	Q45284
4	37	75.5	117	16	Q98218
5	36	73.5	79	2	Q92467
6	36	73.5	243	17	Q81M07
7	36	73.5	285	16	Q97JG0
8	36	73.5	315	16	Q8XX16
9	36	73.5	373	2	Q93SR2
10	36	73.5	373	16	Q910X3
11	36	73.5	458	11	Q9CS06
12	36	73.5	547	11	Q9WVS5
13	36	73.5	914	17	Q8U4Z5
14	35	71.4	51	2	Q86256
15	35	71.4	72	2	Q9AHY8
16	35	71.4	76	16	Q8RHG0

17	35	71.4	116	16	Q8XCF2	Q8XCF2 escherichia
18	35	71.4	154	2	Q8VM90	Q8VM90 rhizobium
19	35	71.4	164	2	Q05140	Q05140 proteus mir
20	35	71.4	164	2	Q8RIY8	Q8RIY8 xenorhabdus
21	35	71.4	164	16	Q9K0U4	Q9K0U4 vibrio chol
22	35	71.4	164	16	Q8ZSC8	Q8ZSC8 yersinia pe
23	35	71.4	164	16	Q8Z815	Q8Z815 salmonella
24	35	71.4	169	16	Q8YCG7	Q8YCG7 brucella me
25	35	71.4	180	16	Q97G14	Q97G14 streptococ
26	35	71.4	519	3	Q9H0J2	Q9H0J2 schizosach
27	34	69.4	43	16	Q9K810	Q9K810 bacillus ha
28	34	69.4	159	16	Q9ABY9	Q9ABY9 caldibacter
29	34	69.4	253	16	Q8Z768	Q8Z768 salmonella
30	34	69.4	283	5	Q8S0W4	Q8S0W4 encyrtallito
31	34	69.4	373	2	Q9A003	Q9A003 pseudomonas
32	34	69.4	373	2	Q9AK03	Q9AK03 pseudomonas
33	34	69.4	373	2	Q91958	Q91958 pseudomonas
34	34	69.4	397	2	Q8U402	Q8U402 melanosarc
35	34	69.4	402	17	Q8U371	Q8U371 melanosarc
36	34	69.4	415	1	Q9V2T2	Q9V2T2 melanosarc
37	34	69.4	415	17	Q8T1Z5	Q8T1Z5 melanosarc
38	34	69.4	481	2	Q50210	Q50210 melittangiu
39	34	69.4	500	10	Q9Y0P0	Q9Y0P0 arabidopsis
40	34	69.4	501	16	Q9P089	Q9P089 ureaplasma
41	34	69.4	502	10	Q22057	Q22057 arabidopsis
42	34	69.4	502	10	Q22576	Q22576 arabidopsis
43	34	69.4	502	10	Q9P0H1	Q9P0H1 arabidopsis
44	34	69.4	502	10	Q24629	Q24629 arabidopsis
45	34	69.4	508	10	Q9SLX4	Q9SLX4 nicotiana t

ALIGNMENTS

RESULT 1
ID Q58439 PRELIMINARY: PRT; 317 AA.
AC Q58439.
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Putative magnesium and cobalt transport protein CORA.
GN MJ1033.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota: Methanococci: Methanococcales;
OC Methanococcaceae: Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / AFCC 47067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bolt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Cocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reick C.L.,
RA Overbeek R., Kirkness E.F., Weisslock K.G., Merrick J.M., Glöck A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utechtack T.K., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii." Science 273:1058-1073(1996).
RL Science 273:1058-1073(1996).
RT -I- FUNCTION: PLAYS A ROLE IN THE TRANSPORT OF MAGNESIUM AND COBALT
IONS (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC EMBL: U67546; AAB99037.1; -.
DR TIGR: MJ1033; -.
DR InterPro: IPR002523; CORA.
DR InterPro: IPR004488; CORA_transp.
DR Pfam: PF01544; CORA; 1.
DR TIGRfam: TIGR00383; CORA; 1.
KW Hypothetical protein; Magnesium; Cobalt; Transmembrane; Transport;
Complete proteome.

FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 291 311 POTENTIAL.
 SQ SEQUENCE 317 AA; 37141 MW; 552A0B378131E0D CRC64;

Query Match
 Best Local Similarity 83.7%; Score 41; DB 17; Length 317;
 Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
 DB 44 LSKRIGLVSVD 54

RESULT 2

Q977W1 PRELIMINARY; PRT; 984 AA.
 AC Q977W1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Topoisomerase V.
 GN TOP5 OR MK1436.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21267361; PubMed=11353838;
 RA Belova G.I., Prasad R., Kozayvkin S.A., Lake J.A., Wilson S.H.,
 RA Slesarev A.I.;
 RT "A type IB topoisomerase with DNA repair activities."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:6015-6020(2001).
 RN 12
 RP SEQUENCE FROM N.A.
 RX STRAIN=AV19 / DSM 6324 / JCM 9639;
 RA MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Mal'yn A.G., Rogozin E.V., Kozayvkin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 and monophyly of archaeal methanogens."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL: AF011944; AAK60014.1;
 DR EMBL: AF010436; AAM02649.1;
 DR InterPro: IPR000445; Hhh.
 DR Pfam: PF00633; HHH.1.
 KW Isomerase; Complete proteome.
 SQ SEQUENCE 984 AA; 112207 MW; 63ANCC517757ID6 CRC64;

Query Match
 Best Local Similarity 79.6%; Score 39; DB 17; Length 984;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
 DB 203 IARLGLSVSE 213

RESULT 3

O45284 PRELIMINARY; PRT; 352 AA.
 AC O45284;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE C31A11.9 protein.
 GN C31A11.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.A.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z83218; CAB05691.1;
 DR InterPro: IPR003002; 7TM_chemol.
 DR InterPro: IPR000168; 7TM_nematode.
 DR Pfam: PF01461; 7tm_4; 1.
 SQ SEQUENCE 352 AA; 40920 MW; A512P50F03860C9 CRC64;

Query Match
 Best Local Similarity 77.6%; Score 38; DB 5; Length 352;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
 DB 170 MSERLGLNVE 180

RESULT 4

O98218 PRELIMINARY; PRT; 117 AA.
 AC O98218;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein mlj9408.
 GN Mlj9408.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN 11
 RP SEQUENCE FROM N.A.
 RX STRAIN=MAFF303099;
 RA MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura Y.,
 RA Kishida Y., Miyokawa C., Kohara M., Matsuno M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimizu S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003015; BAB54438.1;
 DR InterPro: IPR000390; DUF7.
 DR Pfam: PF00893; DUF7.
 KW Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 117 AA; 12444 MW; 1D01801CDB574AF3 CRC64;

Query Match
 Best Local Similarity 75.5%; Score 37; DB 16; Length 117;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 9
 DB 97 LSKRIGLVS 105

RESULT 5

O924G7 PRELIMINARY; PRT; 79 AA.
 AC O924G7;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

```

DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Hypothetical 8.5 kDa protein (CGE011).
GN ORF80 OR CGE011.
OS Escherichia coli, and
OS Salmonella typhimurium.
OG Plasmid pKM101, and Plasmid IncN R46.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 602;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; and S.typhimurium; PLASMID=pKM101, AND INCN R46;
RX MEDLINE=97445167; PubMed=9300052;
RA Delver E.P., Belogurov A.A.;
RT "Organization of the leading region of IncN plasmid pKM101 (R46): a
RT regulation controlled by CUP sequence elements.";
RL J. Mol. Biol. 271:13-30(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; PLASMID=pKM101;
RA Belogurov A.A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 66-79 FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=INCN R46;
RX MEDLINE=93328690; PubMed=8393008;
RA Belogurov A.A., Delver E.P., Rodkevich O.V.;
RT "Plasmid pKM101 encodes two nonhomologous antirestriction proteins
RT (Arda and Ardb) whose expression is controlled by homologous
RT regulatory sequences.";
RL J. Bacteriol. 175:4843-4850(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=INCN R46;
RA Belogurov A.A., Hall R.M., Wihans S.C., Woodgate R.;
RT "Complete sequence of R46 and its deletion derivative, pKM101.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U72482; AAD17386.1;
DR EMBL; AY046276; AAL13409.1;
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 79 AA; 8482 MW; 9EF624EBBD1E297 CRC64;

Query Match
Best Local Similarity 73.5%; Score 36; DB 2; Length 79;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
|:|:|:|:|
DB 69 LAKKIGTISE 79

RESULT 6
Q8TMU7 PRELIMINARY; PRT; 243 AA.
AC Q8TMU7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE Predicted protein.
GN MA2554.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZA / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atencio D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Deaulliano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tiller A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,

```

```

RA Hedderich R., Ingram-Smith C., Kuelner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam I.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Melcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RI Genome Res. 12:532-542(2002).
DR EMBL; AF010950; AAM05936.1;
KW Complete proteome.
SQ SEQUENCE 243 AA; 28216 MW; 0A84089F7430575C CRC64;

Query Match
Best Local Similarity 73.5%; Score 36; DB 17; Length 243;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
|:|:|:|:|
DB 52 LAKKIGTISE 62

RESULT 7
Q97JG0 PRELIMINARY; PRT; 285 AA.
AC Q97JG0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DE Uncharacterized conserved protein, Y11T (B.subtilis) family.
GN CAC1325.
OC Clostridium acetobutylicum.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Bleton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.W., Dubois J., Qiu D., Hitt J., Wolf Y.I.,
RA Tatunov R.L., Sabatle F., Doucelle-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
KI J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AF007644; AAK79294.1;
DR InterPro; IP0007340; IUP0161.
DR Pfam; PF02568; IUP0161; 1.
KW Complete proteome.
SQ SEQUENCE 285 AA; 30797 MW; 51A2742392A0C906 CRC64;

Query Match
Best Local Similarity 73.5%; Score 36; DB 16; Length 285;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
|:|:|:|:|
DB 144 LDKRIGTISE 154

RESULT 8
Q8XX16 PRELIMINARY; PRT; 315 AA.
AC Q8XX16;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Putative transcriptional regulator (Repressor) transcription regulator
DE protein.
GN DALR OR RSC2127 OR RS01506.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

```

OC Ralstonia.
 OC NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Arliguenaue F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brothier P., Camus J.C., Calolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siglier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
 RL Nature 415:497-502(2002).
 DR EMBL; AF646068; CAD15834.1; -
 KM Complete proteome.
 SQ SEQUENCE 315 AA; 33903 MW; 76DDC8460B0866F CRC64;

Query Match 73.5%; Score 36; DB 16; Length 315;
 Best Local Similarity 72.7%; Pred. No. 57;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LS KRI:SVSE 11
 Db 70 LKRG:SVSE 80

RESULT 9

ID Q93SR2 PRELIMINARY; PRT; 373 AA.
 AC Q93SR2;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE Putative muconate cycloisomerase I CatB.
 GN CATB.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-P111;
 RA Totterova S., Armengaud J., Cajthaml T., Demmerova K., Brenner V.,
 RA Pieper D.H.;
 RT "Cloning, isolation and expression of benzoate 1,2-dioxygenase
 RT originated from the strain Pseudomonas putida P111";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY026914; AAK52296.1; -
 DR InterPro: IPR001354; MR_MLE.
 DR Pfam: PF02148; MR_MLE; 1.
 DR Pfam: PF02746; MR_MLE; 1.
 DR PROSITE: PS00908; MR_MLE_1; UNKNOWN_1.
 DR PROSITE: PS00909; MR_MLE_2; UNKNOWN_1.
 KW isomerase.
 SQ SEQUENCE 373 AA; 40131 MW; E7E791D73B8CC6F4 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 373;
 Best Local Similarity 77.8%; Pred. No. 68;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KRI:SVSE 11
 Db 118 KRI:SVSE 126

RESULT 10

ID Q910X3 PRELIMINARY; PRT; 373 AA.
 AC Q910X3;
 DT 01-MAR-2001 (TRENBLREL. 16, Created)
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)

DE Muconate cycloisomerase I.
 GN CATB OR PA2509.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ANCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Watterer P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.L., Gollity L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.K.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004678; AAG05897.1; -
 DR HSSP; P08310; IMC.
 DR InterPro: IPR001354; MR_MLE.
 DR Pfam: PF01188; MR_MLE; 1.
 DR Pfam: PF02746; MR_MLE; 1.
 DR PROSITE: PS00908; MR_MLE_1; 1.
 DR PROSITE: PS00909; MR_MLE_2; 1.
 KM isomerase; Complete proteome.
 SQ SEQUENCE 373 AA; 40045 MW; 949B0FAB3744E76 CRC64;

Query Match 73.5%; Score 36; DB 16; Length 373;
 Best Local Similarity 77.8%; Pred. No. 68;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KRI:SVSE 11
 Db 118 KRI:SVSE 126

RESULT 11

ID Q9CS06 PRELIMINARY; PRT; 458 AA.
 AC Q9CS06;
 DT 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Chaperonin subunit 8 (theta) (fragment).
 GN Cpn8.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Clariata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE=EMBRYONIC LIVER;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
 RA Saito T., Ozaki Y., Gojobori T., Hono H., Kusukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuenli P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli L., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli T., Mombereis P.,
 RA Nordone P., Ring B., Ringwald W., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyo-Oka K., Wang K.H., Wertz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;

RF "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK010882; BAB27244.1; -.
 DR HSSP: P48424; 1A6D.
 DR MGD: MGI:107183; Cct8.
 DR InterPro: IPR002194; Chaperonin_TCP-1.
 DR InterPro: IPR001844; Chaperin_Cpn60.
 DR InterPro: IPR002423; Cpn60/TCP-1.
 DR Pfam: PF00118; Cpn60_TCP1; 1.
 DR PRINTS: PR00298; CHAPERONIN60.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00750; TCP1_1; 1.
 DR PROSITE: PS00751; TCP1_2; 1.
 DR PROSITE: PS00995; TCP1_3; 1.
 DR ATP-binding; Chaperone.
 FT NON_TER 458
 FT 458
 SQ SEQUENCE 458 AA; 49886 MW; B8053A0F4DDC217C CRC64;

Query Match 73.5%; Score 36; DB 11; Length 458;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIGLSVSE 11
 |||||
 Db 107 RIGLSVSE 114

RESULT 12

ID 09WVS5 PRELIMINARY; PRT: 547 AA.
 AC 09WVS5;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Chaperonin containing TCP-1 theta subunit.
 GN COT8 OR COT9.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=128SV;
 RX MEDLINE=92269022; PubMed=10336634;
 RA Kubota H., Yokota S., Yanagi H., Yuta T.;
 RT "Structures and co-regulated expression of the genes encoding mouse
 cytosolic chaperonin COT subunits."
 RL Eur. J. Biochem. 262:492-500(1999).
 DR EMBL: AB022161; BAA81879.1; -.
 DR HSSP: P48424; 1A6D.
 DR MGD: MGI:107183; Cct8.
 DR InterPro: IPR002194; Chaperonin_TCP-1.
 DR InterPro: IPR002423; Cpn60/TCP-1.
 DR Pfam: PF00118; Cpn60_TCP1; 1.
 DR PRINTS: PR00304; TCOMPLEXTCP1.
 DR PROSITE: PS00750; TCP1_1; 1.
 DR PROSITE: PS00751; TCP1_2; 1.
 DR PROSITE: PS00995; TCP1_3; 1.
 DR ATP-binding; Chaperone.
 SQ SEQUENCE 547 AA; 59568 MW; 6BC34987F74B9ACE CRC64;

Query Match 73.5%; Score 36; DB 11; Length 547;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIGLSVSE 11
 |||||
 Db 120 RIGLSVSE 127

RESULT 13

ID 08U425 PRELIMINARY; PRT: 914 AA.
 OY 4 RIGLSVSE 11
 |||||
 Db 120 RIGLSVSE 127

AC 08U425;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alanine-tRNA synthetase.
 GN PF0270.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VCI / DSM 3638 / AFCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AE010151; AAL80394.1; -.
 KW Aminocyl-tRNA synthetase; Complete proteome.
 SQ SEQUENCE 914 AA; 104906 MW; 467F593731BD8BC1 CRC64;

Query Match 73.5%; Score 36; DB 17; Length 914;
 Best Local Similarity 54.5%; Pred. No. 1.9e+02;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSKRIGLSVSE 11
 ::||:|:|
 Db 317 VAKRWGISVEE 327

RESULT 14

ID 086256 PRELIMINARY; PRT: 51 AA.
 AC 086256;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE leucine responsive element (fragment).
 GN LRP.
 OS Klebsiella oxyloca.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OX NCBI_TaxID=571;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Albrecht C., Kleiner D.;
 RT Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
 REGULATORS.
 DR EMBL: Y16963; CAA76565.1; -.
 DR InterPro: IPR000485; ASNC_trans_reg.
 DR PRINTS: PR00033; UTHASNC.
 KW DNA-binding; Transcription regulation.
 FT NON_TER 51
 FT 51
 SQ SEQUENCE 51 AA; 5857 MW; AE2B4C696F788F03 CRC64;

Query Match 71.4%; Score 35; DB 2; Length 51;
 Best Local Similarity 87.5%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSKRIGLS 8
 ||||:|
 Db 34 LSKRWGLS 41

RESULT 15

ID 09AHY8 PRELIMINARY; PRT: 72 AA.
 AC 09AHY8;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE LRP (Fragment).
 GN LRP.

OS Photorhabdus luminescens (Xenorhabdus luminescens).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Photorhabdus.
RN MBL_TaxID=29488;
RP SEQUENCE FROM N.A.
RC STRAIN=NC19;
RX MEDLINE=21225535; PubMed=11325940;
RA Cliche T.A., Bindum S.B., Horswill A.R., Ensign J.C.;
RT "A Phosphopantetheinyl Transferase Homolog Is Essential for
RT Photorhabdus luminescens To Support Growth and Reproduction of the
RT Entomopathogenic Nematode Heterorhabditis bacteriophora";
RL J. Bacteriol. 183:3117-3126(2001).
CC -! SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AF288086; AAK16099.1; -.
DR InterPro; IPR000485; ASNC_trans_reg.
DR Pfam; PF01037; ASNC_trans_reg; 1.
DR PRINTS; PR00033; HTHASNC.
DR SMART; SM00344; HTH_ASNC; 1.
DR PROSITE; PS00519; HTH_ASNC_FAMILY; 1.
DR DNA-binding; Transcription regulation.
KM NON_TER 72 72
FT SEQUENCE 72 AA; 8407 MW; 803932BD8EDB2DB CRC64;
SQ

Query Match 71.4%; Score 35; DB 2; Length 72;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSKRIGLS 8
|||||
Db 34 LSKRVGLS 41

Search completed: January 15, 2003, 12:36:03
Job time : 32.5714 secs

[illegible]

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0133256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142194.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144532.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147028.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 9; DH 21; length 178;
Best Local Similarity 18.2%; Pred. No. 2, 1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 1 LXXXXXXXXX 11
Db 64 LSSSSSSSSSE 74

RESULT 2

AAG49541 standard; Protein; 178 AA.

AC AAG49541;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 62684.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142804.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136762.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138054.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160774.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 9; DB 21; Length 231;
Best Local Similarity 18.2%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXE 11
Db 117 LSSSSSSSSSE 127

RESULT 4
AA049540 ID AA049540 standard; Protein; 231 AA.

AC AA049540;
DI 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 62683.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134296.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145819.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160881.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 50.0%; Score 9; DB 21; Length 231;
Best Local Similarity 18.2%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXX 11
Db 117 LSSSSSSSE 127
RESULT 5
AAG49539
ID AAG49539 standard; Protein; 239 AA.
XX
AC AAG49539;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SWQ ID NO: 62682.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN BP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139495.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0138460.
PR 18-JUN-1999; 99US-0138461.
PR 18-JUN-1999; 99US-0138462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0138817.
PR 22-JUN-1999; 99US-0138899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142520.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 9; DB 21; Length 239;
Best Local Similarity 18.2%; Pred. No. 2.4e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 125 LSSSSSSSSSE 135

RESULT 6
ID ABB69532 standard; Protein; 274 AA.

AC ABB69532:

DE 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 35388.

XX Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

PE 23-MAR-2001; 2001MO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

PI WPI: 2001-656660/75.

DR N-PSDB; ABL13635.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PI interactions -

XX Disclousure ; SEQ ID NO 35388; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins
CC (AB57737-AB572072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 274 AA;

SO Query Match 50.0%; Score 9; DB 22; Length 274;

Best Local Similarity 18.2%; Pred. No. 2.5e+02;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11

DB 66 LTTTSTTSASE 76

RESULT 7
ID AAB58817 standard; Protein; 324 AA.

AC AAB58817:

DT 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 525.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnerrary; anticoagulant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.

OS WO200055173-A1.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

PE 12-MAR-1999; 99US-0124270.

PR (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

PI WPI: 2000-611515/58.

DR N-PSDB; AAF21720.

XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -

PS Claim 11; Page 958-960; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the

CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnerrary; anticoagulant;
CC antibacterial; antifungal; antiparasitic; cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and agonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 324 AA;

SO Query Match 50.0%; Score 9; DB 21; Length 324;

Best Local Similarity 18.2%; Pred. No. 2.7e+02;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11

DB 46 LMAAAMAAAE 56

RESULT 8
ID AAG05407 standard; Protein; 375 AA.

AC AAG05407:

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 1803.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-15; 9; 99US-0121180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0133407.
PR 04-MAY-1999; 99US-0133744.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138547.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0138454.
PR 18-JUN-1999; 99US-0138455.
PR 18-JUN-1999; 99US-0138456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0138462.
PR 18-JUN-1999; 99US-0138463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144652.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148568.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155146.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 9; DB 21; length 375;
Best Local Similarity 18.2%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 60 LSSSSSSWSE 70

RESULT 9
AAG05406
ID AAG05406 standard; Protein: 376 AA.
XX AAG05406;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 1802.
XX
RW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.
XX
PV EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136332.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140383.
PR 23-JUN-1999; 99US-0140384.
PR 24-JUN-1999; 99US-0140655.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

```
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145819.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 9; DB 21; Length 376;
Best Local Similarity 18.2%; Pred. No. 2.9e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Caps 0;

QY 1 LXXXXXXXE 11
Db 61 LSSSSSSTSE 71

RESULT 10
ID AAY21800
AC AAY21800;
DF 20-SEP-1999 (first entry)
DE Human vacuolar ATPase subunit AC45 (HAC45).
XX Vacuolar ATPase subunit AC45;
KW cancer; immune disorder; human.
OS Homo sapiens.
XX US5932444-A.
XX PN 03-AUG-1999.
XX PD
XX PF 28-OCT-1997; 97US-0959011.
XX
```

PR 28-OCT-1997; 97US-0959011.
 XX
 PA (INCYT-) INCYTE PHARM INC.
 XX
 PI Corley NC, Hillman JL, Shah P;
 XX WPI; 1999-443597/37.
 DR N-PSDB; AAK81298.
 XX
 PT Vacuolar ATPase subunit AC45 for treating cancer and immune
 disorders
 XX
 PS Claim 1: Fig 1A-F; 32pp; English.
 CC This represents a human vacuolar ATPase subunit AC45 (HAC45). Vacuolar
 CC ATPase provides most of the energy required for ion transport processes
 CC within the vacuolar system in eukaryotic cells. The polypeptide can be
 CC produced recombinantly by culturing host cells transformed with a vector
 CC comprising the HAC45 nucleic acid. Compositions comprising the HAC45
 CC polynucleotide sequences are useful in the treatment of cancers and
 CC immune disorders. The sequence complementary to the HAC45-encoding
 CC sequence is useful for inhibiting expression of naturally occurring
 CC HAC45.
 CC
 XX
 SQ Sequence 470 AA;
 Query Match 50.0%; Score 9; DB 20; Length 470;
 Best Local Similarity 18.2%; Pred. No. 3.2e+02;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 LXXXXXXXXE 11
 Db 32 LAAAAAAAAAE 42
 RESULT 11
 AAM93671
 ID AAM93671 standard; Protein; 470 AA.
 XX
 AC AAM93671;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3556.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PE 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94606.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8: SEQ ID NO 3556; 1380pp + sequence listing; English.
 CC The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 CC
 XX
 SQ Sequence 470 AA;
 Query Match 50.0%; Score 9; DB 22; Length 470;
 Best Local Similarity 18.2%; Pred. No. 3.2e+02;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 LXXXXXXXXE 11
 Db 32 LAAAAAAAAAE 42
 RESULT 12
 ABB65883
 ID ABB65883 standard; Protein; 803 AA.
 XX
 AC ABB65883;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 24441.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams K, Li PWJ, Myers BW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; AB109986.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 24441; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB161175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 803 AA;

Query Match 50.0%; Score 9; DB 22; Length 803;
 Best Local Similarity 18.2%; Pred. No. 4.1e-02;
 Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXX 11
 DB 3 LTTTTSAAAF 13

RESULT 13
 ABB68397
 ID ABB68397 standard; Protein: 2768 AA.

AC ABB68397;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 31983.

KW Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN W0200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PT 11-JUL-2000; 2000US-0614150.

PS (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL12500.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 31983; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 2768 AA;

Query Match 50.0%; Score 9; DB 22; Length 2768;

Best Local Similarity 18.2%; Pred. No. 7e+02;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXX 11

DB 2558 LSSSTTSSTTE 2568

RESULT 14
 ABB68812
 ID ABB68812 standard; Protein: 99 AA.

XX ABB68812;
 AC 26-MAR-2002 (first entry)
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 33228.

KW Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN W0200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PT 11-JUL-2000; 2000US-0614150.

PS (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL12915.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 33228; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 99 AA;

Query Match 44.4%; Score 8; DB 22; Length 99;

Best Local Similarity 18.2%; Pred. No. 1e+03;

Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXX 11

DB 75 LSSSTTSSTTE 85

RESULT 15
 AAB79413
 ID AAB79413 standard; Protein: 104 AA.

AC AAB79413;

DT 30-APR-2001 (first entry)

DE Corynebacterium glutamicum SWP protein sequence SEQ ID NO:342.

KW Corynebacterium glutamicum; carbon metabolism and energy production; SWP protein; sugar metabolism and oxidative phosphorylation protein; true chemical production; organic acid; proteoglycan amino acid; nonproteogenic amino acid; purine base; pyrimidine base; nucleoside; nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol; carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;

KW diagnosis; Corynebacterium diphtheriae; evolutionary study.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-IB00943.
XX
PR 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031412.
PR 08-JUL-1999; 99DE-1031413.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031424.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031431.
PR 08-JUL-1999; 99DE-1031433.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031562.
PR 08-JUL-1999; 99DE-1031634.
PR 09-JUL-1999; 99DE-1032180.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032230.
PR 09-JUL-1999; 99US-0143208.
PR 14-JUL-1999; 99DE-1032924.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033005.
PR 27-AUG-1999; 99DE-1040765.
PR 31-AUG-1999; 99US-0151572.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042123.
PR 03-SEP-1999; 99DE-1042125.
XX
PA (BADI) BASF AG.
XX
PI Pompeius M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-061975/07.
XX
N-PSDB; AAF71530.
XX
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes -
XX
PS Claim 20: Page 622-624; 1246pp; English.
XX
CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in
CC AAF79243 to AAF 79633 which are involved in carbon metabolism and
CC energy production. The C. glutamicum SMP gene can be used in vectors
CC (II) for expression in host cells and production or modulation of
CC production of fine chemicals, such as, an organic acid, a proteinogenic
CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC (III) encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to
CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC in evolutionary studies, in determining SMP protein regions required
CC for function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH).

XX
SQ Sequence 104 AA:
Query Match 44.4%; Score 8; DB 22; Length 104;
Best Local Similarity 18.2%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
CY 1 LXXXXXXXE 11
DB 83 LATISATTAEE 93

Search completed: January 15, 2003, 11:20:12
Job time : 32.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 : Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec

Title: US-09-823-649a-1
Perfect score: 18
Sequence: 1 LXXXXXXXXXXE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA*
1: /cgn2.6/prodata/1/aa/5A.COMB.pep:*
2: /cgn2.6/prodata/1/aa/5B.COMB.pep:*
3: /cgn2.6/prodata/1/aa/6A.COMB.pep:*
4: /cgn2.6/prodata/1/aa/6B.COMB.pep:*
5: /cgn2.6/prodata/1/aa/PCUS.COMB.pep:*
6: /cgn2.6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	50.0	25	6 5169933-42	Patent No. 5169933
2	50.0	470	2 US-08-958-011-1	Sequence 1, Appl
3	44.4	134	4 US-09-199-637A-243	Sequence 243, App
4	44.4	136	3 US-09-128-155-11	Sequence 11, Appl
5	44.4	167	3 US-09-128-155-7	Sequence 7, Appl
6	44.4	178	3 US-09-128-155-2	Sequence 2, Appl
7	44.4	218	4 US-09-293-625-2	Sequence 2, Appl
8	44.4	917	1 US-08-245-295-2	Sequence 2, Appl
9	44.4	917	1 US-08-481-130-2	Sequence 2, Appl
10	44.4	917	1 US-08-656-984A-2	Sequence 2, Appl
11	44.4	917	1 US-08-485-604-2	Sequence 2, Appl
12	44.4	917	2 US-08-487-595-2	Sequence 2, Appl
13	44.4	2137	4 US-09-134-001C-4463	Sequence 4463, Ap
14	38.9	37	2 US-08-180-524-1	Sequence 1, Appl
15	38.9	37	2 US-08-180-524-8	Sequence 8, Appl
16	38.9	37	2 US-08-180-524-9	Sequence 9, Appl
17	38.9	37	2 US-08-180-524-10	Sequence 10, Appl
18	38.9	37	2 US-08-975-166-1	Sequence 1, Appl
19	38.9	37	2 US-08-975-166-8	Sequence 8, Appl
20	38.9	37	2 US-08-975-166-9	Sequence 9, Appl
21	38.9	37	2 US-08-975-166-10	Sequence 10, Appl
22	38.9	37	4 US-09-117-121-39	Sequence 39, Appl
23	38.9	37	4 US-09-344-529-4	Sequence 4, Appl
24	38.9	109	2 US-08-646-981-6	Sequence 6, Appl
25	38.9	137	1 US-08-258-026A-18	Sequence 18, Appl
26	38.9	137	5 PCT-US95-07541-18	Sequence 18, Appl
27	38.9	142	1 US-08-336-618-18	Sequence 18, Appl

28	38.9	142	5 PCT-US92-03993-7	Sequence 7, Appl
29	38.9	163	4 US-09-562-737-79	Sequence 79, Appl
30	38.9	170	1 US-08-127-499A-20	Sequence 20, Appl
31	38.9	170	1 US-08-482-847-20	Sequence 20, Appl
32	38.9	235	3 US-08-924-747-24	Sequence 24, Appl
33	38.9	235	4 US-09-247-373B-24	Sequence 24, Appl
34	38.9	235	4 US-09-296-715-24	Sequence 24, Appl
35	38.9	252	4 US-09-395-689-5	Sequence 5, Appl
36	38.9	300	3 US-08-705-875A-4	Sequence 4, Appl
37	38.9	300	3 US-08-705-875A-6	Sequence 6, Appl
38	38.9	300	4 US-09-220-731-21	Sequence 21, Appl
39	38.9	328	4 US-09-125-891-2	Sequence 2, Appl
40	38.9	328	4 US-09-125-891-4	Sequence 4, Appl
41	38.9	344	4 US-09-463-702A-42	Sequence 42, Appl
42	38.9	360	4 US-09-231-227-2	Sequence 2, Appl
43	38.9	375	1 US-07-817-917-1	Sequence 1, Appl
44	38.9	411	1 US-08-399-561-2	Sequence 2, Appl
45	38.9	441	2 US-08-491-835-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
5169933-42
Patent No. 5169933
APPLICANT: ANDERSON, DAVID C.; MORGAN, CHARLES JR.; FRITZBERG,
ALAN R.; NICHOLS, EVERETT J.
TITLE OF INVENTION: CAVALENTLY-LINKED COMPLEXES AND METHODS
FOR ENHANCED CYTOTOXICITY AND IMAGING
NUMBER OF SEQUENCES: 45
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/390,241
FILING DATE: 07-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 232,337
FILING DATE: 15-AUG-1988
SEQ ID NO: 42
LENGTH: 25
5169933-42

Query Match 50.0%; Score 9; DB 6; Length 25;
Best Local Similarity 18.2%; Pred. No. 4.6;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 LXXXXXXXXXXE 11
DB 13 LAAAAAAAAAASE 2;
RESULT 2
US-08-959-011-1
Sequence 1, Application US/08959011
Patent No. 5932444
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: VACUOLAR ATPASE SUBUNIT AC45
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/959,011
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0412 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 470 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CRIBNOT01
CLONE: 676592
US-08-959-011-1

Query Match 50.0%; Score 9; DB 2; Length 470;
Best Local Similarity 18.2%; Pred. No. 17;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXXXE 11
DB 32 LAAAAAAAAAE 42

RESULT 3
US-09-199-637A-243
Sequence 243, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Manojan-Miklos, Shalina
APPLICANT: Tao, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Tsoukalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 243
LENGTH: 134
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-243

Query Match 44.4%; Score 8; DB 4; Length 134;
Best Local Similarity 18.2%; Pred. No. 88;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXXXE 11
DB 123 LSSASATTE 133

RESULT 4
US-09-128-155-11

Sequence 11, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 136
TYPE: PRT
ORGANISM: Homo sapiens
US-09-128-155-11

Query Match 44.4%; Score 8; DB 3; Length 136;
Best Local Similarity 18.2%; Pred. No. 89;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXXXE 11
DB 12 LASSLSASAE 22

RESULT 5
US-09-128-155-7
Sequence 7, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 167
TYPE: PRT
ORGANISM: Homo sapiens
US-09-128-155-7

Query Match 44.4%; Score 8; DB 3; Length 167;
Best Local Similarity 18.2%; Pred. No. 97;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXXXE 11
DB 43 LASSLSASAE 53

RESULT 6
US-09-128-155-2
Sequence 2, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03

EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 2
LENGTH: 178
TYPE: PRT
ORGANISM: Homo sapiens
US-09-128-155-2

Query Match 44.4%; Score 8; DB 3; Length 178;
Best Local Similarity 18.2%; Pred. No. 1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 54 LASSLSASAE 64

RESULT 7
US-09-293-625-2
Sequence 2, Application US/09293625
Patent No. 6342371
GENERAL INFORMATION:
APPLICANT: Young, Peter R.
APPLICANT: McDonnell, Peter C
TITLE OF INVENTION: INTERLEUKIN-1 HOMOLOGUE, IL-1H4
FILE REFERENCE: GP-70607
CURRENT APPLICATION NUMBER: US/09/293,625
CURRENT FILING DATE: 1999-04-16
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentl Ver. 2.0
SEQ ID NO: 2
LENGTH: 218
TYPE: PRT
ORGANISM: Homo sapiens
US-09-293-625-2

Query Match 44.4%; Score 8; DB 4; Length 218;
Best Local Similarity 18.2%; Pred. No. 1.1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 94 LASSLSASAE 104

RESULT 8
US-08-245-295-2
Sequence 2, Application US/08245295
Patent No. 5700658
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, Suite 6300
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentl Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,295
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27866/32055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-245-295-2

Query Match 44.4%; Score 8; DB 1; Length 917;
Best Local Similarity 18.2%; Pred. No. 2.1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXXX 11
DB 282 LVATAVAVASE 292

RESULT 9
US-08-481-130-2
Sequence 2, Application US/08481130
Patent No. 5702917
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentl Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,130
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32713
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-130-2

Query Match 44.4% Score 8; DB 1; Length 917;
Best Local Similarity 18.2%; Pred. No. 2.1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXX 11
Db 282 LVATATATASE 292

RESULT 10
US-08-656-984A-2
Sequence 2, Application US/08656984A
Patent No. 5753502
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,984A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,604
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33321
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-656-984A-2

Query Match 44.4% Score 8; DB 1; Length 917;
Best Local Similarity 18.2%; Pred. No. 2.1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 LXXXXXXXXX 11
Db 282 LVATATATASE 292

RESULT 11
US-08-485-604-2
Sequence 2, Application US/08485604
Patent No. 5773293
GENERAL INFORMATION:
APPLICANT: W.P. W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,604
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
APPLICATION NUMBER: US 08/102,852

FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32715
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-604-2

Query Match 44.4%; Score 8; DB 1; Length 917;
Best Local Similarity 18.2%; Pred. No. 2, 1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 282 LVATATATASE 292

RESULT 12
US-08-487-595-2
Sequence 2, Application US/08487595
Patent No. 5852170
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,595
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:

NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32714
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-487-595-2

Query Match 44.4%; Score 8; DB 2; Length 917;
Best Local Similarity 18.2%; Pred. No. 2, 1e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 282 LVATATATASE 292

RESULT 13
US-09-134-001C-4463
Sequence 4463, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4463
LENGTH: 2137
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4463

Query Match 44.4%; Score 8; DB 4; Length 2137;
Best Local Similarity 18.2%; Pred. No. 3e+02;
Matches 2; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 LXXXXXXXE 11
DB 1701 LSDSTSTSE 1711

RESULT 14
US-08-180-524-1
Sequence 1, Application US/08180524
Patent No. 5848537
GENERAL INFORMATION:
APPLICANT: Tripp, Matthew
APPLICANT: Lusk, Lance
APPLICANT: Rhodes, Thomas
APPLICANT: Huijge, Nick
APPLICANT: Kot, Edward
APPLICANT: Chicoye, Etzer
APPLICANT: Barney, Michael C.
APPLICANT: Bower, Patricia A.
APPLICANT: Cronan, Charles L.
TITLE OF INVENTION: METHOD OF EXPRESSING ANTIFREEZE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

Search completed: January 16, 2003, 11:04:51
Job time : 1662.14 secs

```
* 40724 40823: gap of unknown length
* 40824 42680: contig of 1857 bp in length
* 42681 42780: gap of unknown length
* 42781 43816: contig of 1036 bp in length
* 43817 43916: gap of unknown length
* 43917 45637: contig of 1721 bp in length
* 45638 45737: gap of unknown length
* 45738 47557: contig of 1820 bp in length
* 47558 47657: gap of unknown length
* 47658 50323: contig of 2666 bp in length
* 50324 50423: gap of unknown length
* 50424 52036: contig of 1613 bp in length
* 52037 52136: gap of unknown length
* 52137 54072: contig of 1936 bp in length
* 54073 54172: gap of unknown length
* 54173 56825: contig of 2653 bp in length
* 56826 56925: gap of unknown length
* 56926 60336: contig of 3411 bp in length
* 60337 60436: gap of unknown length
* 60437 62006: contig of 1570 bp in length
* 62007 62106: gap of unknown length
* 62107 64744: contig of 2638 bp in length
* 64745 64844: gap of unknown length
* 64845 68236: contig of 3392 bp in length
* 68237 68336: gap of unknown length
* 68337 71193: contig of 2857 bp in length
* 71194 71293: gap of unknown length
* 71294 73460: contig of 2167 bp in length
* 73461 73560: gap of unknown length
* 73561 76278: contig of 2718 bp in length
* 76279 76378: gap of unknown length
* 76379 78639: contig of 2261 bp in length
* 78640 78739: gap of unknown length
* 78740 82062: contig of 3323 bp in length
* 82063 82162: gap of unknown length
* 82163 84918: contig of 2756 bp in length
* 84919 85018: gap of unknown length
* 85019 88476: contig of 3458 bp in length
* 88477 88576: gap of unknown length
* 88577 91511: contig of 2935 bp in length
* 91512 91611: gap of unknown length
* 91612 93532: contig of 3741 bp in length
* 93533 95452: gap of unknown length
* 95453 98310: contig of 2858 bp in length
* 98311 98410: gap of unknown length
* 98411 101300: contig of 2890 bp in length
* 101301 101400: gap of unknown length
* 101401 103847: contig of 4447 bp in length
* 103848 105947: gap of unknown length
* 105948 109639: contig of 3692 bp in length
* 109640 109739: gap of unknown length
* 109740 113323: contig of 3584 bp in length
* 113324 113423: gap of unknown length
* 113424 116402: contig of 2979 bp in length
* 116403 116502: gap of unknown length
* 116503 119250: contig of 2748 bp in length
* 119251 119350: gap of unknown length
* 119351 123808: contig of 4458 bp in length
```

Alignment Scores:

```
Pred. No.: 7.27e+03 Length: 180674
Score: 40.00 Matches: 8
Percent Similarity: 90.91% Conservative: 2
Best Local Similarity: 72.73% Mismatches: 1
Query Match: 81.63% Indels: 0
DB: 2 Gaps: 0
```

US-09-823-649a-6 (1-11) x AC094431 (1-180674)

QY 1 Leuserlysaqllleqlyleuserlysaqlu 11

DB 117650 TTAAAGAGAGATTGGACTTACACTTCGAG 117618

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:02:09 ; Search time 28.1429 seconds

(without alignments)
52.083 Million cell updates/sec

Title: US-09-823-649a-6
Perfect score: 49
Sequence: 1 LSKRIGLSVSE 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	49	100.0	11	23	AA847796	Native DNA polymerase
2	49	100.0	11	23	AA848268	Native DNA polymerase
3	49	100.0	609	13	AA823174	Mutant thermophilic
4	49	100.0	690	13	AA823173	Mutant thermophilic
5	49	100.0	754	13	AA823172	Mutant thermophilic
6	49	100.0	800	13	AA823171	Mutant thermophilic
7	49	100.0	856	13	AA823170	Mutant thermophilic
8	49	100.0	892	13	AA823169	Mutant thermophilic
9	49	100.0	892	13	AA823168	Taf DNA polymerase
10	36	73.5	128	21	AA85382	Human 5' EST relat

11	36	73.5	258	21	AA843412	Human cancer assoc
12	36	73.5	548	16	AA879585	Mouse CBT-beta su
13	36	73.5	548	20	AA87062	Renal cancer assoc
14	36	73.5	555	22	ABG21964	Novel human diagno
15	36	73.5	584	22	ABG21966	Novel human diagno
16	35	71.4	90	19	AA85848	S. pneumoniae deri
17	35	71.4	1043	22	ABG08827	Novel human diagno
18	34	69.4	79	20	AA812988	Human secreted pro
19	34	69.4	373	23	AA823653	Stabilizing and de
20	34	69.4	382	21	AA823939	Arabidopsis thalia
21	34	69.4	382	21	AA843430	Arabidopsis thalia
22	34	69.4	408	20	AA860505	Human normal blad
23	34	69.4	502	21	AA823938	Arabidopsis thalia
24	34	69.4	502	21	AA843429	Arabidopsis thalia
25	34	69.4	502	23	AA891449	Arabidopsis thalia
26	34	69.4	532	21	AA823937	Herbicideally activ
27	34	69.4	532	21	AA843428	Arabidopsis thalia
28	34	69.4	779	21	AA894997	Human secreted pro
29	33	67.3	59	23	AA806956	Human OREF protein
30	33	67.3	62	18	AA820505	H. pylori cytoplas
31	33	67.3	79	20	AA807970	Human secreted pro
32	33	67.3	165	18	AA820621	H. pylori cytoplas
33	33	67.3	549	23	AA829511	Streptococcus poly
34	33	67.3	822	23	AA807951	S. salivarius 1,6-
35	33	67.3	889	22	AA820268	Human lung specif
36	33	67.3	956	22	AA893299	Human protein sequ
37	33	67.3	1046	20	AA828106	Amino acid sequenc
38	32	65.3	37	22	AA807530	Human polypeptide
39	32	65.3	91	23	AA805858	Human OREF protein
40	32	65.3	119	22	AA811177	Bacillus anthracis
41	32	65.3	119	22	AA849471	Listeria monocytog
42	32	65.3	182	20	AA838894	Neisseria meningit
43	32	65.3	256	20	AA838895	Neisseria meningit
44	32	65.3	256	20	AA838896	Neisseria meningit
45	32	65.3	311	22	AA808750	Human melanocortin

ALIGNMENTS

RESULT 1
AA847796 standard; peptide: 11 AA.
XX
AC AA847796;
XX
DP 25-MAR-2002 (first entry)
XX
DE Native DNA polymerase mol1 #6.
XX
KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KW Transverse transcription reaction; fluorescence; cyanine; thermocative;
KW dye; amplification.
XX
OS Thermophilic africanus.
XX
PN EP152062-A2.
XX
PD 07-NOV-2001.
XX
PF 12-APR-2001; 2001EP-0109341.
XX
PR 18-APR-2000; 2000US-198336P.
XX
PA (HOFF) HOFFMANN IA ROCHE & CO AG F.
XX
PI Smith ES, Elstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenrunner NJ, Wang AM;
XX
WP1: 2002-076891/11.
XX
PT Reverse transcribing an RNA, comprises performing a reverse
transcriptase polymerase chain reaction amplification of a mixture

PT using a mutant thermocactive DNA polymerase -
 XX
 PS Claim 2; Page 4; 23pp; English.
 XX
 CC The sequences given in AAB47791-97 represent the native forms of motifs
 CC derived from DNA polymerases used in the method of the invention. The
 CC method for reverse transcribing an RNA, comprises treating a transverse
 CC transcription reaction mixture which comprises the RNA, a primer, a
 CC divalent cation, and a mutant thermocactive DNA polymerase. These motifs
 CC represent a conserved region which affects the DNA polymerase's ability
 CC to incorporate dideoxynucleotides labelled with fluorescein and cyanine
 CC family dyes. The method of the invention is useful in reverse
 CC transcription/amplification reactions. The method provides improved
 CC reverse transcription efficiency at lower enzyme concentrations relative
 CC to previous high temperature reverse transcription methods, and provide
 CC a wider range of usable salt concentrations. The use of the mutant DNA
 CC polymerase provides faster reverse transcription extension rates, and
 CC consequently less time is needed for the reaction.
 XX
 SQ Sequence 11 AA;
 XX
 Query Match 100.0%; Score 49; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LSKRIGLVSVE 11
 |||||
 DB 1 LSKRIGLVSVE 11
 |||||
 RESULT 2
 AAM48268
 ID AAM48268 standard; Peptide: 11 AA.
 XX
 AC AAM48268;
 XX
 DT 25-MAR-2002 (first entry)
 XX
 DE Native DNA polymerase motif #17.
 XX
 KM DNA polymerase; reverse transcription; primer; divalent cation; mutant;
 KM transverse transcription reaction; fluorescein; cyanine; thermocactive;
 KM dye; amplification.
 XX
 OS Thermosipho africanus.
 XX
 PN EP1152062-A2.
 XX
 PD 07-NOV-2001.
 XX
 PF 12-APR-2001; 2001EP-0109341.
 XX
 PR 18-APR-2000; 2000US-198336P.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Smith ES, Elfstrom CM, Gelland DH, Higuchi RG, Myers TW;
 PI Schoenbrunner NJ, Wang AM;
 XX
 DR WP1: 2002-076891/11.
 XX
 PT Reverse transcribing an RNA, comprises performing a reverse
 PT transcriptase polymerase chain reaction amplification of a mixture
 PT using a mutant thermocactive DNA polymerase -
 XX
 PS Disclosure; Page 7; 23pp; English.
 XX
 CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
 CC native forms of motifs derived from DNA polymerases used in the method
 CC of the invention. The method for reverse transcribing an RNA, comprises
 CC treating a transverse transcription reaction mixture which comprises the
 CC RNA, a primer, a divalent cation, and a mutant thermocactive DNA
 CC polymerase. These motifs represent a conserved region which affects the

CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
 CC fluorescein and cyanine family dyes. The method of the invention is
 CC useful in reverse transcription/amplification reactions. The method
 CC provides improved reverse transcription efficiency at lower enzyme
 CC concentrations relative to previous high temperature reverse
 CC transcription methods, and provide a wider range of usable salt
 CC concentrations. The use of the mutant DNA polymerase provides faster
 CC reverse transcription extension rates, and consequently less time is
 CC needed for the reaction.
 XX
 SQ Sequence 11 AA;
 XX
 Query Match 100.0%; Score 49; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LSKRIGLVSVE 11
 |||||
 DB 1 LSKRIGLVSVE 11
 |||||
 RESULT 3
 AAR23174
 ID AAR23174 standard; Protein: 609 AA.
 XX
 AC AAR23174;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase enzyme MET-ILE 285 TAF.
 XX
 KM 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
 XX
 OS Thermosipho africanus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..2
 FT note="residues 2-284 deleted from the native
 FT sequence"
 XX
 PN W09206200-A.
 XX
 PD 16-APR-1992.
 XX
 PF 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETUS) CETUS CORP.
 XX
 PI Abramson RD, Gelland DH;
 XX
 DR WP1: 1992-150885/18.
 DR N-PSDB: AAQ24334.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC The sequence is that of a mutant of Thermosipho africanus polymerase
 CC mutant MET-ILE 284 TAF, having a different amt. of 5'-3' exonuclease
 CC activity than the native enzyme. Thermostable DNA polymerases
 CC are useful in many recombinant DNA techniques, esp. nucleic acid
 CC amplification by PCR, self-sustained sequence replication (SSR)
 CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
 CC may facilitate higher sensitivity allelic discrimination in a combined
 CC polymerase ligation chain reaction (PLCR) assay. An enhanced amt. of
 CC 5'-3' exonuclease activity may be desirable in enzymes used in
 CC homogeneous assays for the amplification and detection of a target

CC nucleic acid sequence. Mutation of the DNA encoding particular
 CC regions of the enzymes can be used to prepare a range of recombinant
 CC proteins having 5'-3' exonuclease activity to a complete lack of
 CC activity.
 CC See also AAR23140-79 and AAR23722.
 CC
 CC Sequence 609 AA;
 CC
 CC

Query Match 100.0%; Score 49; DB 13; Length 609;
 Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Gaps 0;

OY 1 LSKRIGLSVSE 11
 |||||
 DB 457 LSKRIGLSVSE 467

RESULT 4
 AAR23173
 ID AAR23173 standard; Protein; 690 AA.
 XX
 AC AAR23173;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase enzyme MFT-THR 204 TAF.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
 XX
 OS Thermosipho africanus.
 XX

Key Location/Qualifiers
 FH Misc-difference 1..2 - "residues 2-203 deleted from the native
 FT sequence"
 FT
 PN W09206200-A.
 XX
 PD 16-APR-1992.
 XX
 PE 30-SEP-1991; 91MO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelfand DH;
 XX
 DR WPI; 1992-150885/18.
 DR N-PSDB; AAQ24333.
 XX
 XX

PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC

The sequence is that of a mutant of Thermosipho africanus polymerase
 mutant MFT-THR 204 TAF, having a different amt. of 5'-3' exonuclease
 CC activity than the native enzyme. Thermostable DNA polymerases
 CC are useful in many recombinant DNA techniques, esp. nucleic acid
 CC amplification by PCR, self-sustained sequence replication (SSR)
 CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
 CC may facilitate higher sensitivity allelic discrimination in a combined
 CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
 CC 5'-3' exonuclease activity may be desirable in enzymes used in
 CC homogeneous assays for the amplification and detection of a target
 CC nucleic acid sequence. Mutation of the DNA encoding particular
 CC regions of the enzymes can be used to prepare a range of recombinant
 CC proteins having 5'-3' exonuclease activity to a complete lack of
 CC activity.

CC See also AAR23140-79 and AAR23722.
 CC
 CC Sequence 690 AA;
 CC
 CC

Query Match 100.0%; Score 49; DB 13; Length 690;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSKRIGLSVSE 11
 |||||
 DB 538 LSKRIGLSVSE 548

RESULT 5
 AAR23172
 ID AAR23172 standard; Protein; 754 AA.
 XX
 AC AAR23172;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase enzyme MFT-GLU 140 TAF.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
 XX
 OS Thermosipho africanus.
 XX

Key Location/Qualifiers
 FH Misc-difference 1..2 /note= "residues 2-139 deleted from the native
 FT sequence"
 FT
 PN W09206200-A.
 XX
 PD 16-APR-1992.
 XX
 PE 30-SEP-1991; 91MO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelfand DH;
 XX
 DR WPI; 1992-150885/18.
 DR N-PSDB; AAQ24332.
 XX
 XX

PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC

The sequence is that of a mutant of Thermosipho africanus polymerase
 mutant designated MFT-GLU 140 TAF, having a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. However the residue
 CC present in the given sequence at position 140 is an Asp.
 CC Thermostable DNA polymerases are useful in many recombinant DNA
 CC techniques, esp. nucleic acid amplification by PCR, self-sustained
 CC sequence replication (SSR) and high temp. DNA sequencing. The absence
 CC of 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be
 CC desirable in enzymes used in homogeneous assays for the amplification
 CC and detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare a
 CC range of recombinant proteins having 5'-3' exonuclease activity to a
 CC complete lack of activity.
 CC See also AAR23140-79 and AAR23722.
 CC
 CC Sequence 754 AA;
 CC
 CC


```

RESULT 8
AAR23169
ID AAR23169 standard; Protein; 892 AA.
XX
XX AAR23169.
AC
XX
XX 22-OCT-1992 (first entry)
DT
XX
XX Mutant thermostable DNA polymerase enzyme Asp37 Taf.
DE
XX
XX 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR.
KM
XX
XX Thermosipho africanus.
OS
XX
XX Key Location/Qualifiers
FT Misc-difference 37 /note="Gly in native sequence"
FT
XX
XX W09206200-A.
PN
XX
XX 16-APR-1992.
PD
XX
XX 30-SEP-1991; 91WO-US07035.
PF
XX
XX 28-SEP-1990; 90US-0590213.
PR
XX
XX 28-SEP-1990; 90US-0590466.
PR
XX
XX 28-SEP-1990; 90US-0590490.
PA
XX
XX (CERTU ) CERTUS CORP.
PI
XX
XX Abramson RD, Gelfand DH.
PI
XX
XX WPI: 1992-150885/18.
DR
XX
XX N-PSDB; AAQ24329.
XX
XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
XX
XX Claim 11; Page 59; 185pp; English.
XX
XX The sequence is that of a mutant of Thermosipho africanus polymerase
CC mutant Asp37 Taf, having a different amt. of 5'-3' exonuclease
CC activity than the native enzyme. Thermostable DNA polymerases
CC are useful in many recombinant DNA techniques, esp. nucleic acid
CC amplification by PCR, self-sustained sequence replication (SSR)
CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
CC may facilitate higher sensitivity allele discrimination in a combined
CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
CC 5'-3' exonuclease activity may be desirable in enzymes used in
CC homogeneous assays for the amplification and detection of a target
CC nucleic acid sequence. Mutation of the DNA encoding particular
CC regions of the enzymes can be used to prepare a range of recombinant
CC proteins having 5'-3' exonuclease activity to a complete lack of
CC activity.
CC See also AAR23140-79 and AAR23722.
CC
XX
XX Sequence 892 AA;
SQ

```

```

XX
XX AAR23122;
AC
XX
XX 27-OCT-1992 (first entry)
DT
XX
XX Taf DNA polymerase I.
DE
XX
XX Thermostability; PCR; polymerase chain reaction;
KM reverse transcription; exonuclease; proofreading.
KM
XX
XX Thermosipho africanus.
OS
XX
XX W09206202-A.
PN
XX
XX 16-APR-1992.
PD
XX
XX 26-SEP-1991; 91WO-US07076.
PF
XX
XX 28-SEP-1990; 90US-0590490.
PR
XX
XX (CERTU ) CERTUS CORP.
PA
XX
XX Abramson RD, Gelfand DH, Greenfield L, Lawyer FC, Reichert FL;
PI
XX
XX WPI: 1992-150887/18.
DR
XX
XX N-PSDB; AAQ23917.
XX
XX Thermostable DNA polymerase from Thermosipho africanus - prep'd
PT by purification from cells or by expression of Taf polymerase gene
PT in host cells
XX
XX Claim 7; Page 68; 80pp; English.
XX
XX The thermostable polymerase from Thermosipho africanus (Taf) has
CC DNA polymerase activity, reverse transcriptase activity, and opt.
CC 5'-3' and/or 3'-5' exonuclease activity. The Taf enzyme does not
CC become irreversibly denatured for relatively short exposures to
CC temperatures of 90-100 deg.C. The Taf enzyme has an optimum
CC temperature above 45 deg.C and exhibits activity over a broad
CC temperature range of 37-90 deg.C. The Taf polymerase is suitable
CC for use in PCR and in other techniques such as DNA sequencing,
CC nick-translation and reverse transcription.
CC See also AAQ23918-Q23961.
CC
XX
XX Sequence 892 AA;
SQ

```

```

Query Match 100.0%; Score 49; DB 13; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSKRIGLSYSE 11
| | | | | | | | | |
DB 740 LSKRIGLSYSE 750

```

```

RESULT 10
AAV65382
ID AAV65382 standard; Protein; 128 AA.
XX
XX AAV65382.
AC
XX
XX 01-FEB-2000 (first entry)
DT
XX
XX Human 5' EST related polypeptide SEQ ID NO:1543.
DE
XX
XX Human: 5' EST: expressed sequence tag; secreted protein; diagnosis;
KM gene therapy; chromosome mapping; upstream regulatory sequence;
KM forensic; location; development; protein synthesis; stability;
KM regulation; identification.
XX
XX Homo sapiens.
OS
XX
XX W09953051-A2.
PN

```

```

XX PD 21-OCT-1999.
XX PF 09-APR-1999. 99MO-IB00712.
XX PR 09-APR-1998. 98US-0057719.
XX PR 28-APR-1998. 98US-0069047.
XX PA (GEST) GENSET.
XX PI Dumas Milne Edwards J, Duclet A, Giordano J;
XX DR WPI: 2000-038446/03.
XX DR N-PSDB: AA42396.
XX PT Novel secreted protein 5' expressed sequence tag sequences used in
XX PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX PS Claim 3; Page 815; 837pp; English.
CC AA42265. o AA43075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AA46651 to
CC AA6338 represent the EST-related proteins corresponding to AA42265 to
CC AA43052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA42249 to AA42264 and AA46644 to AA46650 represent
CC sequences used in the exemplification of the present invention.
XX SQ Sequence 128 AA;
Query Match 73.5%; Score 36; DB 21; Length 128;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RIGLSVSE 11
Db 120 RIGLSVSE 127
RESULT 11
AAB43412
ID AAB43412 standard; Protein; 258 AA.
XX AC AAB43412;
XX DT 08-FEB-2001 (first entry)
XX DE Human cancer associated protein sequence SPQ ID NO:857.
XX KW Human: cancer associated gene; cancer antigen; detection; cancer;
XX KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
XX KW antidiabetic; antitumor; antihemorrhagic; antithrombotic; antiviral;
XX KW antidiabetic; antihypertensive; antidiabetic; antidiabetic; antidiabetic;
XX KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX KW vasotropic; antipsoriatic; antianagenic; gene therapy; inflammation;
XX KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
XX KW allergic reaction; graft versus host disease; organ rejection;
XX KW haemostatic; thrombolytic; cardiovascular disorder; infection;
XX KW neurological disease; drug screening.

```

```

OS Homo sapiens.
XX PN WO20005350-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05882.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI: 2000-587533/55.
XX DR N-PSDB: AAC77621.
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
XX PT useful for treating or diagnosing e.g. cancer -
XX PS Claim 11; Page 1411-1412; 2352pp; English.
CC AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antitumor; antihemorrhagic; antithrombotic; antiviral;
CC antidiabetic; antihypertensive; antidiabetic; antidiabetic; antidiabetic;
CC dermatological; neuroprotective; thrombolytic; coagulant;
CC nootropic; vasotropic; antipsoriatic; antidiabetic; antidiabetic;
CC polynucleotides and polypeptides can be used for preventing, treating or
CC polymers, medical conditions and diagnosing pathological conditions.
CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.
XX SQ Sequence 258 AA;
Query Match 73.5%; Score 36; DB 21; Length 258;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RIGLSVSE 11
Db 148 RIGLSVSE 155
RESULT 12
AAR79585
ID AAR79585 standard; Protein; 548 AA.
XX AC AAR79585;
XX DT 17-JUL-1996 (first entry)
XX DE Mouse CCT-theta subunit sequence, deduced from clone pTbeta1.
XX KW Mouse CCT-theta subunit sequence, deduced from clone pTbeta1.
XX KW Chaperonin containing TCP-1; CCT; subunit; protein folding complex;
XX KW murine; denaturation; renaturation.
XX OS Mus musculus.
XX PN WO9520654-A1.
XX PD 03-AUG-1995.

```

XX 31-JAN-1995; 95WO-GB00192.
 PF 09-SEP-1994; 94GB-0018234.
 XX 31-JAN-1994; 94GB-0001791.
 PR (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.
 PA Ashworth A, Kubota H, Willison KR;
 PI WPI: 1995-275439/36.
 DR N-PSDB; AAG97801.
 XX
 PF Foiding protein complex sub-unit(s) - with ability to form complex
 in vitro, useful in facilitating folding of proteins
 PS Claim 19; Fig 4; 94pp; English.

CC Mouse cDNAs encoding novel TCP-1-related proteins were isolated
 CC using a combination of methods. Clones pDeltat2 and pZeta12 were
 CC isolated by cross-hybridisation with C.elegans cDNA probes recovered
 CC from a 5'-expressed sequence tag collection. The clones pDeltat2,
 CC pZeta12 and pDeltat1 were isolated by cross-hybridisation with
 CC human pr¹es recovered by PCR of human HT1080 cell line cDNA using
 CC degenerate primers from a conserved region of TCP-1 and TF55. Clone
 CC pDeltat2 was isolated during a mouse testis cDNA sequencing project
 CC and clone pZeta12 was recovered by hybridisation with a mouse cDNA
 CC PCR product which was made with primers derived from the sequence of
 CC a TCP-1 related gene fragment recovered accidentally from a human
 CC kidney cDNA library during a screen for ion transport channel
 CC genes. TCP-1 and the proteins encoded by the cDNA clones are
 CC individual subunits of a heteromeric chaperonin which has been
 CC designated "chaperonin, containing TCP-1" (CCT). The chaperonin acts
 CC as a protein folding complex. The present sequence is that of the
 CC CCT-theta subunit.

Sequence 548 AA;

Query Match 73.5%; Score 36; DB 16; Length 548;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIGLSVSE 11
 DB 120 RIGLSVSE 127

RESULT 13

AA07062
 ID AA07062 standard; Protein: 548 AA.

AC AA07062;

DT 02-JUL-1999 (first entry)

DE Renal cancer associated antigen precursor sequence.

KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer.

OS Homo sapiens.

PN WO9904265-A2.

PD 28-JAN-1999.

PF 15-JUL-1998; 98WO-US14679.

PR 22-JUN-1998; 98US-0102322.

PR 17-JUL-1997; 97US-0896164.

PR 10-OCT-1997; 97US-0061599.

PR 10-OCT-1997; 97US-0061765.

PR 10-OCT-1997; 97US-0948705.
 PR 11-OCT-1997; 97GB-0021697.

XX (LUDWIG) INST CANCER RES.

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
 PI Pfeundschnub M, Sahin U, Scanlan MJ, Stockert E;
 PI Tureci O;

DR WPI: 1999-132448/11.

PT New isolated cancer associated nucleic acids and polypeptides -
 PT isolated using sera from cancer patients; used to develop products
 PT for the diagnosis, monitoring or treatment of cancers

PS Disclosure; Page 458-459; 787pp; English.

CC The invention relates to a method for diagnosing a disorder characterised
 CC by expression of a human cancer associated antigen precursor coded for by
 CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
 CC biological sample isolated from a subject with an agent that specifically
 CC binds to the NAM, an expression product or a fragment of an expression
 CC product complexed with an HLA molecule; and (b) determining the
 CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer.

Sequence 548 AA;

Query Match 73.5%; Score 36; DB 20; Length 548;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIGLSVSE 11
 DB 120 RIGLSVSE 127

RESULT 14

ABG21964
 ID ABG21964 standard; Protein: 555 AA.

AC ABG21964;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #21955.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Dymanac RT, Liu C, Tang YF;

DR WPI: 2001-639362/73.
 N-PSDB; AAS86151.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 52323; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging or sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. Abg00010-Abg30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 555 AA;
 Query Match 73.5%; Score 36; DB 22; Length 555;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 RIGLSYSE 11
 Db 127 RIGLSYSE 134
 RESULT 15
 ABG21966
 ID ABG21966 standard; Protein; 584 AA.
 XX
 AC ABG21966;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21957.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensics;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS86153.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 52325; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging or sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. Abg00010-Abg30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 584 AA;
 Query Match 73.5%; Score 36; DB 22; Length 584;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 RIGLSYSE 11
 Db 108 RIGLSYSE 115
 Search completed: January 15, 2003, 11:20:21
 Job time : 32.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 (Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec)

Title: US-09-823-649A-6
Perfect score: 49
Sequence: 1 LSKRIGLSVSE 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/Backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	49	100.0	892 1 US-07-977-434-12	Sequence 12, Appl
2	49	100.0	892 1 US-08-458-819-12	Sequence 12, Appl
3	46	100.0	892 5 PCT-US91-07035-12	Sequence 12, Appl
4	36	73.5	548 4 US-08-687-590-31	Sequence 31, Appl
5	34	69.4	79 4 US-08-905-223-272	Sequence 272, App
6	34	69.4	373 4 US-09-586-875-1	Sequence 1, Appl
7	33	104.6	4 US-09-199-637A-13	Sequence 13, Appl
8	32	65.3	332 2 US-08-662-560-2	Sequence 2, Appl
9	32	65.3	332 2 US-08-780-749A-6	Sequence 6, Appl
10	32	65.3	332 3 US-08-706-281A-16	Sequence 16, Appl
11	32	65.3	332 4 US-09-097-231-16	Sequence 16, Appl
12	32	65.3	332 4 US-08-870-511-6	Sequence 6, Appl
13	32	65.3	332 4 US-08-870-511-8	Sequence 8, Appl
14	32	65.3	332 4 US-08-870-511-10	Sequence 10, Appl
15	32	65.3	332 4 US-08-870-511-12	Sequence 12, Appl
16	32	65.3	370 2 US-08-729-214-27	Sequence 27, Appl
17	32	65.3	370 2 US-08-928-934-27	Sequence 27, Appl
18	32	65.3	373 3 US-09-028-934-32	Sequence 32, Appl
19	32	65.3	15281 2 US-08-471-119A-2	Sequence 2, Appl
20	31	63.3	485 6 5320958-2	Patent No. 5320958
21	31	63.3	785 4 US-09-134-001C-3866	Sequence 3866, Ap
22	30	61.2	20 2 US-08-564-972-20	Sequence 20, Appl
23	30	61.2	169 4 US-09-516-914-7	Sequence 7, Appl
24	30	61.2	193 2 US-08-564-972-8	Sequence 8, Appl
25	30	61.2	194 1 US-08-148-058A-27	Sequence 27, Appl
26	30	61.2	194 1 US-08-148-058A-29	Sequence 29, Appl
27	30	61.2	194 1 US-08-478-042-27	Sequence 27, Appl

28	30	61.2	194 1 US-08-478-042-29	Sequence 29, Appl
29	30	61.2	194 2 US-08-645-215-27	Sequence 27, Appl
30	30	61.2	194 2 US-08-645-215-29	Sequence 29, Appl
31	30	61.2	194 2 US-08-466-604-27	Sequence 27, Appl
32	30	61.2	194 2 US-08-466-604-29	Sequence 29, Appl
33	30	61.2	216 2 US-08-737-716-4	Sequence 4, Appl
34	30	61.2	220 4 US-09-134-001C-5073	Sequence 5073, Ap
35	30	61.2	251 4 US-09-627-376-13	Sequence 13, Appl
36	30	61.2	437 2 US-08-737-716-2	Sequence 2, Appl
37	30	61.2	524 4 US-09-186-276B-56	Sequence 56, Appl
38	30	61.2	524 4 US-08-842-445-56	Sequence 56, Appl
39	30	61.2	524 4 US-09-186-188B-56	Sequence 56, Appl
40	30	61.2	535 2 US-08-564-972-1	Sequence 1, Appl
41	30	61.2	535 4 US-09-171-969-9	Sequence 9, Appl
42	30	61.2	535 4 US-09-250-131-2	Sequence 2, Appl
43	30	61.2	678 1 US-08-844-085-2	Sequence 2, Appl
44	30	61.2	836 1 US-08-216-971-2	Sequence 2, Appl
45	30	61.2	836 2 US-08-812-979-2	Sequence 2, Appl

ALIGNMENTS

RESURF 1
US-07-977-434-12
Sequence 12, Application US/07977434
Patent No. 5465591
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
NUMBER OF INVENTIONS: THERMOSTABLE DNA POLYMERASES
CORRESPONDENCE ADDRESS:
ADDRESS: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: 7
SOFTWARE: Wordperfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28 SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28 SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28 SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseert
REGISTRATION NUMBER: 31,822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 892 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-977-434-12

Query Match 100.0%; Score 49; DB 1; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLSYSE 11 --
|||||
Db 740 LSKRIGLSYSE 750

RESULT 2
US-08-458-819-12
Sequence 12, Application US/08458819
Patent No. 5795762
GENERAL INFORMATION:
APPLICANT: Geifand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: WordPerfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,819
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,434
FILING DATE: 23-FEB-1993
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseert
REGISTRATION NUMBER: 31,822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 892 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-819-12

Query Match 100.0%; Score 49; DB 1; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLSYSE 11
|||||
Db 740 LSKRIGLSYSE 750

RESULT 3
PCT-US91-07035-12
Sequence 12, Application PC/TUS9107035
GENERAL INFORMATION:
APPLICANT: Geifand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Celus Corporation
STREET: 1400 Fifty-third Street
CITY: Emeryville
STATE: California
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.0

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US91/07035
;; FILING DATE: 19910930
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 590,490
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 590,466
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 590,213
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 523,394
;; FILING DATE: 15-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 143,441
;; FILING DATE: 12-JAN-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 063,509
;; FILING DATE: 17-JUN-1987
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 899,241
;; FILING DATE: 22-AUG-1986
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 746,121
;; FILING DATE: 15-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US90/07641
;; FILING DATE: 21-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 585,471
;; FILING DATE: 20-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 455,611
;; FILING DATE: 22-DEC-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 609,157
;; FILING DATE: 02-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 557,517
;; FILING DATE: 24-JUL-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sias Ph.D., Stacey R.
;; REGISTRATION NUMBER: 32,630
;; REFERENCE/DOCKET NUMBER: Case No. 2580
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-420-3500
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 892 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; PCT-US91-07035-12

Query Match 100.0%; Score 49; DB 5; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLVSVE 11
|||||
DB 740 LSKRIGLVSVE 750

RESULT 4
US-08-687-590-31
; Sequence 31, Application US/08687590
; Patent No. 6255070
; GENERAL INFORMATION:
; APPLICANT: Willison, Keith Robert
; APPLICANT: Kudota, Hiroshi

;; APPLICANT: Ashworth, Alan
;; TITLE OF INVENTION: Folding Proteins
;; NUMBER OF SEQUENCES: 74
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, Eighth Floor
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States of America
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/687,590
;; FILING DATE: 31-JUL-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/GB95/00192
;; FILING DATE: 31-JAN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9401791.0
;; FILING DATE: 31-JAN-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9418234.2
;; FILING DATE: 09-SEP-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bastian, Kevin L.
;; REGISTRATION NUMBER: 34,774
;; REFERENCE/DOCKET NUMBER: 084619-000000US
;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 548 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-687-590-31

Query Match 73.5%; Score 36; DB 4; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RIGLSVSW 11
|||||
DB 120 RIGLSVSW 127

RESULT 5
US-08-905-223-272
; Sequence 272, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
;; NUMBER OF SEQUENCES: 503
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Knobbe, Martens, Olson & Bear
;; STREET: 501 West Broadway
;; CITY: San Diego
;; STATE: California
;; COUNTRY: USA
;; ZIP: 92101-3505
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy Disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: Win95
;; SOFTWARE: Word
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/905,223

;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Israelson, Ned A.
;; REGISTRATION NUMBER: 29,655
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 235-8550
;; TELEFAX: (619) 235-0176
;; INFORMATION FOR SEQ ID NO: 272:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 79 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE: PROTEIN
;; ORIGINAL SOURCE:
;; ORGANISM: Homo Sapiens
;; TISSUE TYPE: Brain
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: -26...-1
;; IDENTIFICATION METHOD: Von Heljne matrix
;; OTHER INFORMATION: score 10.8
US-08-905-223-272

Query Match 69.4%; Score 34; DB 4; Length 79;
Best Local Similarity 63.6%; Pred. No. 2.4;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSKRIGLSVSE 11
|:|:|:|:|:|
DB 48 LAKTLGLEVSE 58

RESULT 6
US-09-586-875-1
; Sequence 1, Application US/09586875
; Patent No. 6385546
; GENERAL INFORMATION:
; APPLICANT: K... Peter C.
; APPLICANT: G. Gnan, Adrian
; APPLICANT: Helin, Sari
; TITLE OF INVENTION: STABILIZING AND DESTABILIZING PROTEINS
; FILE REFERENCE: 08036-018002
; CURRENT APPLICATION NUMBER: US/09/586,875
; CURRENT FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: 08/971,004
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: 60/030,926
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-09-586-875-1

Query Match 69.4%; Score 34; DB 4; Length 373;
Best Local Similarity 77.8%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 KRIGLSVSE 11
|:|:|:|:|:|
DB 118 KRLGLPYSE 126

RESULT 7
US-09-199-637A-13
; Sequence 13, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:

;; APPLICANT: Ausubel, Frederick
;; APPLICANT: Goodman, Howard M.
;; APPLICANT: Rahme, Laurence G.
;; APPLICANT: Mahajan-Wiklos, Shalina
;; APPLICANT: Tan, Man-Wah
;; APPLICANT: Cao, Hui
;; APPLICANT: Drenkard, Eliana
;; APPLICANT: Tsongalis, John
;; TITLE OF INVENTION: VIRULENCE ASSOCIATED NUCLEIC ACID
;; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
;; FILE REFERENCE: 00786/361002
;; CURRENT APPLICATION NUMBER: US/09/199,637A
;; CURRENT FILING DATE: 1998-11-25
;; PRIOR APPLICATION NUMBER: 60/066,517
;; PRIOR FILING DATE: 1997-11-25
;; NUMBER OF SEQ ID NOS: 437
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13
;; LENGTH: 1046
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-13

Query Match 67.3%; Score 33; DB 4; Length 1046;
Best Local Similarity 77.8%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSKRIGLSV 9
|:|:|:|:|:|
DB 656 LAKRRLSV 664

RESULT 8
US-08-662-560-2
; Sequence 2, Application US/08662560
; Patent No. 5908603
; GENERAL INFORMATION:
; APPLICANT: Lee, Frank
; APPLICANT: Huszar, Dennis
; APPLICANT: Wei, Gu
; TITLE OF INVENTION: SCREENING METHODS FOR COMPOUNDS
; TITLE OF INVENTION: USEFUL IN THE REGULATION OF BODY WEIGHT
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,560
; FILING DATE: 10-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-060
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-662-560-2

Query Match 65.3%; Score 32; DB 2; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 KRIGLAYS 10
11:1:1:1
DB 164 KRIGLAYS 171

RESULT 9
US-08-780-749A-6
Sequence 6, Application US/08780749A
Patent No. 5932779

GENERAL INFORMATION:
APPLICANT: Lee, Frank
APPLICANT: Huszar, Dennis
APPLICANT: Gu, Wei
TITLE OF INVENTION: SCREENING METHODS FOR COMPOUNDS
TITLE OF INVENTION: USEFUL IN THE REGULATION OF BODY WEIGHT
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/780,749A
FILING DATE: 08-JAN-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Laura A. Cortuzzi
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-064
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-780-749A-6

Query Match 65.3%; Score 32; DB 2; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 KRIGLAYS 10
11:1:1:1
DB 164 KRIGLAYS 171

RESULT 10
US-08-706-281A-16
Sequence 16, Application US/08706281A
Patent No. 6100048

GENERAL INFORMATION:
APPLICANT: Cone, Roger D
APPLICANT: Fan, Wei
APPLICANT: Boston, Bruce A
APPLICANT: Kesterton, Robert A
APPLICANT: Lu, Dongxi
TITLE OF INVENTION: Methods and Reagents for Discovering and
TITLE OF INVENTION: Using Mammalian Melanocortin Receptor Agonists and Antagonists
TITLE OF INVENTION: To Modulate Feeding Behavior in Animals
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: vaterlin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,281A
FILING DATE: 04-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6100048nat, Kevin B
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 96,886
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-706-281A-16

Query Match 65.3%; Score 32; DB 3; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 KRIGLAYS 10
11:1:1:1
DB 164 KRIGLAYS 171

RESULT 11
US-09-097-231-16
Sequence 16, Application US/09097231
Patent No. 6278038
GENERAL INFORMATION:
APPLICANT: Cone, Roger D
APPLICANT: Chen, Wenbiao
APPLICANT: Low, Malcolm J
TITLE OF INVENTION: Mammalian Melanocortin Receptor and Uses
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,231
; FILING DATE: 12-Jun-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6278038nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 96,886-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-097-231-16

Query Match          65.3%; Score 32; DB 4; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 KRIGSVS 10
       11:1:1:1
Db      164 KRIGSIS 171

RESULT 12
US-08-870-511-6
; Sequence 6, Application US/08870511
; Patent No. 6287763
; GENERAL INFORMATION:
; APPLICANT: Lee, Frank
; APPLICANT: Huszar, Dennis
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: SCREENING METHODS FOR COMPOUNDS USEFUL IN THE
; TITLE OF INVENTION: REGULATION OF BODY WEIGHT
; FILE REFERENCE: 7853-083
; CURRENT APPLICATION NUMBER: US/08/870,511
; CURRENT FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-870-511-6

Query Match          65.3%; Score 32; DB 4; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 KRIGSVS 10
       11:1:1:1
Db      164 KRIGSIS 171

RESULT 13
US-08-870-511-8
; Sequence 8, Application US/08870511
; Patent No. 6287763
; GENERAL INFORMATION:
; APPLICANT: Lee, Frank
; APPLICANT: Huszar, Dennis
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: SCREENING METHODS FOR COMPOUNDS USEFUL IN THE
; TITLE OF INVENTION: REGULATION OF BODY WEIGHT
; FILE REFERENCE: 7853-083
; CURRENT APPLICATION NUMBER: US/08/870,511
```

```

; CURRENT FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-870-511-8

Query Match          65.3%; Score 32; DB 4; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 KRIGSVS 10
       11:1:1:1
Db      164 KRIGSIS 171

RESULT 14
US-08-870-511-10
; Sequence 10, Application US/08870511
; Patent No. 6287763
; GENERAL INFORMATION:
; APPLICANT: Lee, Frank
; APPLICANT: Huszar, Dennis
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: SCREENING METHODS FOR COMPOUNDS USEFUL IN THE
; TITLE OF INVENTION: REGULATION OF BODY WEIGHT
; FILE REFERENCE: 7853-083
; CURRENT APPLICATION NUMBER: US/08/870,511
; CURRENT FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-870-511-10

Query Match          65.3%; Score 32; DB 4; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 KRIGSVS 10
       11:1:1:1
Db      164 KRIGSIS 171

RESULT 15
US-08-870-511-12
; Sequence 12, Application US/08870511
; Patent No. 6287763
; GENERAL INFORMATION:
; APPLICANT: Lee, Frank
; APPLICANT: Huszar, Dennis
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: SCREENING METHODS FOR COMPOUNDS USEFUL IN THE
; TITLE OF INVENTION: REGULATION OF BODY WEIGHT
; FILE REFERENCE: 7853-083
; CURRENT APPLICATION NUMBER: US/08/870,511
; CURRENT FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-870-511-12

Query Match          65.3%; Score 32; DB 4; Length 332;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

QY 3 KRIGLWS 10
11:1:1:1
Db 164 KRWGISIS 171

Search completed: January 15, 2003, 12:38:25
Job time : 10.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:20:34 ; Search time 10.8571 Seconds
(without alignments)
20.140 Million cell updates/sec

Title: US-09-823-649a-6
Perfect score: 49
Sequence: 1 LSKRIGLSYSE 11

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	11	10	US-09-823-649a-6
2	49	100.0	11	10	US-09-823-649a-17
3	36	73.5	258	10	US-09-925-301-857
4	34	69.4	485	8	US-08-808-031A-2
5	33	67.3	822	10	US-09-981-900B-9
6	31	63.3	112	10	US-09-214-474A-6
7	31	63.3	146	10	US-09-815-242-10980
8	31	63.3	193	10	US-09-888-911-4
9	31	63.3	201	10	US-09-815-242-10879
10	31	63.3	203	10	US-09-815-242-10879
11	31	63.3	310	9	US-10-027-806-44
12	31	63.3	310	9	US-10-034-623-44
13	31	63.3	480	8	US-08-808-031A-42
14	31	63.3	526	9	US-09-738-626-6454
15	31	63.3	565	9	US-09-999-248-4
16	31	63.3	719	10	US-09-815-242-5520
17	31	63.3	782	10	US-09-815-242-12367
18	31	63.3	782	10	US-09-815-242-12984
19	30	61.2	139	10	US-09-867-550-482

20	30	61.2	154	10	US-09-864-761-47352	Sequence 47352, A
21	30	61.2	169	10	US-09-870-162A-7	Sequence 7, Appl1
22	30	61.2	200	9	US-09-908-931B-11	Sequence 11, Appl1
23	30	61.2	200	10	US-09-815-242-5404	Sequence 5404, Ap
24	30	61.2	200	10	US-09-815-242-12578	Sequence 12578, A
25	30	61.2	200	10	US-09-815-242-12778	Sequence 12778, A
26	30	61.2	211	9	US-09-791-932-97	Sequence 97, Appl1
27	30	61.2	251	12	US-10-047-676A-13	Sequence 13, Appl1
28	30	61.2	318	9	US-09-905-291A-341	Sequence 341, Appl
29	30	61.2	318	9	US-09-902-853-341	Sequence 341, Appl
30	30	61.2	318	9	US-09-907-824-341	Sequence 341, Appl
31	30	61.2	318	9	US-09-907-841-341	Sequence 341, Appl
32	30	61.2	318	9	US-09-904-011-341	Sequence 341, Appl
33	30	61.2	318	9	US-10-028-072-374	Sequence 374, Appl
34	30	61.2	318	10	US-09-909-320-341	Sequence 341, Appl
35	30	61.2	318	10	US-09-909-088B-341	Sequence 341, Appl
36	30	61.2	323	9	US-09-791-932-119	Sequence 119, Appl
37	30	61.2	400	9	US-09-738-626-4488	Sequence 4488, Ap
38	30	61.2	1460	9	US-09-568-756-2	Sequence 4488, Ap
39	30	61.2	2383	10	US-09-912-020-302	Sequence 302, Appl
40	29	59.2	11	10	US-09-823-649A-16	Sequence 16, Appl
41	29	59.2	78	9	US-10-001-876-139	Sequence 139, Appl
42	29	59.2	112	9	US-10-051-643-78	Sequence 78, Appl
43	29	59.2	112	9	US-09-880-505-78	Sequence 78, Appl
44	29	59.2	126	9	US-09-738-626-5824	Sequence 5824, Ap
45	29	59.2	144	10	US-09-815-242-5458	Sequence 5458, Ap

ALIGNMENTS

```
RESULT 1
US-09-823-649a-6
: Sequence 6, Appl1 US/09823649A
: Patent No. US20020012970A1
:
: GENERAL INFORMATION:
: APPLICANT: Smith, Edward
: APPLICANT: Elstiom, Carila
: APPLICANT: Gelfand, David
: APPLICANT: Higuchi, Russell
: APPLICANT: Myers, Thomas
: APPLICANT: Schoendranner, Nancy
: APPLICANT: Wang, Alice
:
: TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYME
: FILE REFERENCE: RPA1006
: CURRENT APPLICATION NUMBER: US/09/823,649A
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/198,336
: PRIOR FILING DATE: 2000-04-18
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: Patent In version 3.0
: SEQ ID NO 6
: LENGTH: 11
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: sequence motif
US-09-823-649a-6
Query Match 100.0%; Score 49; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 LSKRIGLSYSE 11
Db 1 LSKRIGLSYSE 11
RESULT 2
US-09-823-649a-17
: Sequence 17, Appl1 US/09823649A
: Patent No. US20020012970A1
: GENERAL INFORMATION:
```

APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
FILE REFERENCE: P01006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 11
TYPE: PRT
ORGANISM: Thermophilus africanus
US-09-823-649A-17

Query Match
Best Local Similarity 100.0%; Score 49; DB 10; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SKRIGLVSSE 11
DB 1 SKRIGLVSSE 11

RESULT 3
US-09-925-301-857
Sequence 857, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: P0106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 857
LENGTH: 258
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-301-857

Query Match
Best Local Similarity 100.0%; Score 36; DB 10; Length 258;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RIGLVSSE 11
DB 148 RIGLVSSE 155

RESULT 4
US-08-808-031A-2
Sequence 2, Application US/08808031A
Patent No. US2002004802A1
GENERAL INFORMATION:
APPLICANT: Inouye, Sumiko
APPLICANT: Hsu, Wei-Yin
APPLICANT: Eagle, Susan
APPLICANT: Inouye, Masayori
TITLE OF INVENTION: PROKARYOTIC REVERSE TRANSCRIPTASE
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:

ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,031A
FILING DATE: 03-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 377(913).5888P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-031A-2

Query Match
Best Local Similarity 69.4%; Score 34; DB 8; Length 485;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKALGLSVSK 165
DB 155 LKALGLSVSK 165

RESULT 5
US-09-981-900B-9
Sequence 9, Application US/09981900B
Patent No. US2002013878A1
GENERAL INFORMATION:
APPLICANT: Sticklen, Masomeh B
APPLICANT: Magbool, Shahina B
APPLICANT: Dale, Bruce E
TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING LIGNINASE AND CELLULOSE WHICH DEG
FILE REFERENCE: MSU 4.1-539
CURRENT APPLICATION NUMBER: US/09/981,900B
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 60/242,408
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 822
TYPE: PRT
ORGANISM: Streptococcus salivarius
US-09-981-900B-9

Query Match
Best Local Similarity 70.0%; Score 33; DB 10; Length 822;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SKRIGLVSSE 11
DB 63 SKRIGLVSSE 72

RESULT 6
US-09-214-474A-6

```
/ Sequence 6, Application US/09214474A
/ Patent No. US20020049311A1
/ GENERAL INFORMATION:
/ APPLICANT: Gress, Michael J.
/ APPLICANT: Hegg, Lisa A.
/ APPLICANT: Li, Hu
/ APPLICANT: Prescott, Catherine D.
/ APPLICANT: Chappell, Amy M.
/ TITLE OF INVENTION: No. US20020049311A1el Rnase P
/ FILE REFERENCE: GMS0036
/ CURRENT APPLICATION NUMBER: US/09/214,474A
/ CURRENT FILING DATE: 1999-01-21
/ PRIOR APPLICATION NUMBER: US 60/057,520
/ PRIOR FILING DATE: 1997-09-04
/ PRIOR APPLICATION NUMBER: PCT/US98/18291
/ PRIOR FILING DATE: 1988-09-03
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6
/ LENGTH: 112
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-214-474A-6
```

```
Query Match 63.3%; Score 31; DB 10; Length 112;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 RIGLVSSE 11
1:|||||:
DB 41 RVLGLSVK 48
```

```
RESULT 7
US-09-815-242-10980
/ Sequence 10980, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10980
/ LENGTH: 146
/ TYPE: PRT
/ ORGANISM: Haemophilus influenzae
US-09-815-242-10980
```

```
Query Match 63.3%; Score 31; DB 10; Length 146;
Best Local Similarity 45.5%; Pred. No. 29;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LSKRIGLVSSE 11
1:|||||:
DB 8 LKKKVGKIKITE 18
```

```
RESULT 8
US-09-888-911-4
/ Sequence 4, Application US/09888911
/ Patent No. US20020111307A1
/ GENERAL INFORMATION:
/ APPLICANT: Glucksmann, Maria Alexandra.
/ APPLICANT: Rudolph-owen, Laura A.
/ TITLE OF INVENTION: 46508, A NOVEL HUMAN PEPTIDYL-TRNA
/ TITLE OF INVENTION: HYDROLASE FAMILY MEMBER AND USES THEREOF
/ FILE REFERENCE: 10448-070001
/ CURRENT APPLICATION NUMBER: US/09/888,911
/ CURRENT FILING DATE: 2001-06-25
/ PRIOR APPLICATION NUMBER: 60/213,688
/ PRIOR FILING DATE: 2000-06-23
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 193
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: consensus sequence
US-09-888-911-4
```

```
Query Match 63.3%; Score 31; DB 10; Length 193;
Best Local Similarity 54.5%; Pred. No. 40;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 LSKRIGLVSSE 11
1:|||||:
DB 30 LSKRIGLVSSE 40
```

```
RESULT 9
US-09-815-242-5030
/ Sequence 5030, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
```


PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5030
LENGTH: 201
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-5030

Query Match 63.3%; Score 31; DB 10; Length 201;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSKRIGLSYS 10
Db 11 VSRRLGISLS 20

RESULT 10
US-09-815-242-10879
Sequence 10879, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10879
LENGTH: 203
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10879

Query Match 63.3%; Score 31; DB 10; Length 203;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSKRIGLSYS 10
Db 11 VSRRLGISLS 20

RESULT 11
US-10-027-806-44
Sequence 44, Application US/10027806
Patent No. US20020160476A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.

APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCOF.002A
CURRENT APPLICATION NUMBER: US/10/027,806
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 44
LENGTH: 310
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-10-027-806-44

Query Match 63.3%; Score 31; DB 9; Length 310;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KRIGLSVSE 11
Db 255 KRIGLKVTD 263

RESULT 12
US-10-034-623-44
Sequence 44, Application US/10034623
Publication No. US20020198365A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCOF.002A
CURRENT APPLICATION NUMBER: US/10/034,623
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/408,020
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: 60/102,294
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 44
LENGTH: 310
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-10-034-623-44

Query Match 64.7%; Score 41; DB 9; Length 310;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KRIGLSVSE 11
Db 255 KRIGLKVTD 263

RESULT 13
US-08-808-031A-42
Sequence 42, Application US/08808031A
Patent No. US20020048902A1
GENERAL INFORMATION:
APPLICANT: Inouye, Sumiko
APPLICANT: Hsu, Mei-Yin
APPLICANT: Bagle, Susan
APPLICANT: Inouye, Masayori
TITLE OF INVENTION: PROKARYOTIC REVERSE TRANSCRIPTASE
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia

STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,031A
FILING DATE: 03-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 377(913).5888P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-031A-42

Query Match 63.3%; Score 31; DB 8; Length 480;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKRIGLSVS 10
Db 153 LAKALGISVS 162

RESULT 14
US-09-738-626-6454
Sequence 6454, Application US/09738626
Publication No. 7520020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAMA, SATOSHI
APPLICANT: WIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6454
LENGTH: 526
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6454

Query Match 63.3%; Score 31; DB 9; Length 526;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
Db 373 LAKVIGLSVPE 383

RESULT 15
US-09-999-248-4
Sequence 4, Application US/09999248
Patent No. US20020176852A1
GENERAL INFORMATION:
APPLICANT: Lambeth, J. David
APPLICANT: Cheng, Guangjie
TITLE OF INVENTION: Mitogenic Oxygenase Regulators
FILE REFERENCE: 05501-0180 43150-266489
CURRENT APPLICATION NUMBER: US/09/999,248
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/249,305
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/251,364
PRIOR FILING DATE: 2000-12-05
PRIOR APPLICATION NUMBER: US 60/289,172
PRIOR FILING DATE: 2001-05-07
PRIOR APPLICATION NUMBER: US 60/289,537
PRIOR FILING DATE: 2001-05-07
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 565
TYPE: PRT
ORGANISM: Homo sapiens
US-09-999-248-4

Query Match 63.3%; Score 31; DB 9; Length 565;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSKRIGLSVS 10
Db 232 LERKIGLSVS 241

Search completed: January 15, 2003, 12:39:52
Job time: 11.8571 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:15:24 ; Search time 12.1429 Seconds

(without alignments)
87.086 Million cell updates/sec

Title: US-09-823-649a-6

Sequence: 1 LSKRIGLSYSE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR:73:*
2: PIR:1:*
3: PIR:2:*
4: PIR:3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	83.7	317	1 H64428	magnesium and coba
2	38	77.6	352	2 T19599	hypothetical prote
3	37	75.5	162	2 E64539	shikimate kinase
4	36	73.5	285	2 C97063	uncharacterized co
5	36	73.5	373	2 D83331	muconate cyclisom
6	36	73.5	548	2 JG4073	TCF-1 containing c
7	35	71.4	116	2 C86120	probable growth in
8	35	71.4	116	2 D49339	cell growth regula
9	35	71.4	116	2 C91279	probable growth in
10	35	71.4	157	2 F69592	azlBCD operon tran
11	35	71.4	164	1 R8RC1R	leucine-responsive
12	35	71.4	164	2 S59991	transcription regu
13	35	71.4	164	2 S59993	transcription regu
14	35	71.4	164	2 S59992	transcription regu
15	35	71.4	164	2 B82142	leucine-responsive
16	35	71.4	164	2 A10167	leucine-responsive
17	35	71.4	164	2 F90750	leucine-responsive
18	35	71.4	164	2 D85614	leucine-responsive
19	35	71.4	164	2 AE0611	leucine-responsive
20	35	71.4	164	2 AB3580	proline dehydrogen
21	35	71.4	180	2 A95138	hypothetical prote
22	34	69.4	43	2 F83915	hypothetical prote
23	34	69.4	159	2 H87258	transcription regu
24	34	69.4	200	1 A37146	ribosomal protein
25	34	69.4	201	1 S65998	conserved hypothet
26	34	69.4	253	2 A10665	probable periplasm
27	34	69.4	375	1 A28630	muconate cyclisom
28	34	69.4	464	2 C69732	PBSX prophage ORF
29	34	69.4	485	1 RRYC62	RNA-directed DNA p

30	34	69.4	500	2 T52642	transcription init
31	34	69.4	501	2 G82896	probable ABC subst
32	34	69.4	502	2 H96671	RNA polymerase sig
33	34	69.4	530	2 E82491	sigma-54 dependent
34	34	69.4	530	2 E82491	acetylcholinestera
35	33	67.3	77	2 B95003	hypothetical prote
36	33	67.3	79	2 A69460	hypothetical prote
37	33	67.3	87	2 H87348	hypothetical prote
38	33	67.3	112	2 F87875	hypothetical prote
39	33	67.3	129	2 F84707	hypothetical prote
40	33	67.3	151	2 C90048	conserved hypothet
41	33	67.3	160	2 D84023	hypothetical prote
42	33	67.3	162	2 C71968	shikimate kinase I
43	33	67.3	175	2 C97138	hypothetical prote
44	33	67.3	263	2 H95407	hypothetical prote
45	33	67.3	302	2 APT033	hypothetical prote

ALIGNMENTS

RESULT 1

H64428 magnesium and cobalt transport protein homolog - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: H64428

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak

; Reich, C.I.; Overbeek, R.; Kirtness, E.P.; Weissbrock, K.G.; Merrick, J.M.; Glodex,

Ison, J.D.; Sadow, P.M.; Hanna, M.C.; Colton, M.D.; Roberts, K.M.; Hursel, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese

A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasch

A:Reference number: A64300; PMID:96337999; PMID:868087

A:Accession: H64428

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-317 <BDL>

A:Cross-references: GB:U67546; GB:L77117; NID:g1591687; PIDN:AA899037.1; PIDN:g1499876

C:Genetics:

A:Map position: F08965089-966042

C:Superfamily: magnesium and cobalt transport protein

Query Match 83.7%; Score 41; DB 1; Length 317;

Best local Similarity 72.7%; Pred. No. 1.8;

Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLSYSE 11

DB 44 LSKRIGLSYSD 14

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Best Local Similarity 63.6%; Pred. No. 8.2;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
:|:|:|:|:|
Db 170 MSERIGLNVSE 180

RESULT 3

E64539
shikimate acid kinase I - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 18-Jun-1999

C:Accession: E64539

R:Romb, J.F.; White, O.; Kerlavang, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, R.; Dodson, P.; Khakhria, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wattey, L.; Nature 388, 539-547, 1997

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: E64539

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-162 <TOM>

A:Cross-references: GB:AE000536; GB:AE000511; NID:92313230; PIDN:AA07220.1; PID:9231323

C:Superfamily: shikimate kinase; shikimate kinase homology

P:1-149/Domain: shikimate kinase homology <SKI>

Query Match 75.5%; Score 37; DB 2; Length 162;
Best Local Similarity 63.6%; Pred. No. 5.9;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
:|:|:|:|:|
Db 36 ISERIGLSVSE 46

RESULT 4

C97063
uncharacterized conserved protein, yitC (B. subtilis) family CAC1325 [imported] - Clostr

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: C97063

R:Noiing, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C97063

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-285 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79294.1; PID:915024255; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1325

C:Superfamily: conserved hypothetical protein yitC

Query Match 73.5%; Score 36; DB 2; Length 285;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
:|:|:|:|:|
Db 144 LDKRIGLSVSE 154

RESULT 5

D83331
mucronate cycloisomerase I PA2509 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: D83331
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; L.

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: D83331

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-373 <STO>

A:Cross-references: GB:AE004678; GB:AE004091; NID:99946560; PIDN:AA605897.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: catB; PA2509

C:Superfamily: mucronate cycloisomerase

Query Match 73.5%; Score 36; DB 2; Length 373;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KRIGLSVSE 11
||:|:|:|:|
Db 118 KRIGLSVSE 126

RESULT 6

JC4073
TCP-1 containing cytosolic chaperonin (CCP) theta chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 20-Aug-1999

C:Accession: JC4073; PC4022; S52867

R:Kubota, H.; Hynes, G.; Williston, K.

Gene 154, 231-236, 1995

A:Title: The eighth Cct gene, Cct9, encoding the theta subunit of the cytosolic chape

A:Reference number: JC4073; MUID:95197008; PMID:7890169

A:Accession: JC4073

A:Molecule type: mRNA

A:Residues: 1-548 <KUR>

A:Cross-references: EMBL:Z37164; NID:9695624; PIDN:CA85521.1; PID:9695625

A:Molecule type: protein

A:Residues: 64-72; 441-450 <KUR>

C:Comment: This protein is involved in folding, transport and assembly of newly synth

C:Genetics:

A:Gene: Cct9

C:Superfamily: molecular chaperone; t-complex-type

C:Keywords: cytosol; molecular chaperone

Query Match 73.5%; Score 36; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RIGLSVSE 11
|||:|:|:|:|
Db 120 RIGLSVSE 127

RESULT 7

C86120
probable growth inhibitor [similarity] - Escherichia coli (strain O157:H7, substrain

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Nov-2001

C:Accession: C86120

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

hler, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoudis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: C86120

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <STO>

A:Cross-references: GB:AE005174; NID:q12519227; PIDN:AA659423.1; GSPDB:GN00145; UMGCP:Z58
A:Experimental source: strain 0157:H7, substrain EDL933
C:Genetics:
A:Gene: chpB

Query Match 71.4%; Score 35; DB 2; Length 116;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
|:|||||: |
Db 92 LAKRIGLADE 102

RESULT 8
D49339
cell growth regulatory protein ChpB - Escherichia coli (strain K-12)

N:Alternate names: pemK-like protein 2

C:Species: Escherichia coli

C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 01-Mar-2002

C:Accession: D49339; S56451; D65234

R:Masuda, Y.; Miyakawa, K.; Nishimura, Y.; Ohtsubo, E.

J. Bacteriol. 175: 6850-6856, 1993

A:Title: chpB and chpB, Escherichia coli chromosomal homologs of the pem locus responsible for growth regulation in Escherichia coli

A:Reference number: A49339; MOID:94042847; PMID:8226627

A:Accession: D49339

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <MAS>

A:Cross-references: GB:D16451; NID:q452513; PIDN:BA03920.1; PID:d100436; PID:g506196

R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.

Nucleic Acids Res. 23, 2105-2119, 1995

A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.

A:Reference number: S56314; MOID:9534362; PMID:7610040

A:Accession: S56451

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-116 <BURL>

A:Cross-references: EMBL:U14003; NID:q1263172; PIDN:AA97122.1; PID:g537067

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.

A:Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MOID:97426617; PMID:9278503

A:Accession: D65234

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-116 <BLAT>

A:Cross-references: GB:AE000494; GB:U00096; NID:q1790670; PIDN:AACT7182.1; PID:q1790672

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: chpB

C:Keywords: DNA binding

Query Match 71.4%; Score 35; DB 2; Length 116;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
|:|||||: |
Db 92 LAKRIGLADE 102

RESULT 9
C91279
probable growth inhibitor [imported] - Escherichia coli (strain 0157:H7, substrain R1MD

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C:Accession: C91279

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and 9
A:Reference number: A99629; MOID:21156231; PMID:11258796
C:Genetics:
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA038626.1; PID:q13364680; GSPDB:GN00154
A:Experimental source: strain 0157:H7, substrain R1MD 050952
C:Genetics:
A:Gene: hcs5203

Query Match 71.4%; Score 35; DB 2; Length 116;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 11
|:|||||: |
Db 92 LAKRIGLADE 102

RESULT 10

azlBCD operon transcription repressor azlB - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000

C:Accession: F69592; T44776

R:Kunst, F.; Brouillet, S.; Moszer, J.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Ber

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.;

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Filtz, C.; Fujita, Y.; Fuma, S.; Galiczi, A.; Gal

lechi, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Iardino

Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardino

A:Authors: Lamber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y.M.; Ogawa, K.; Ogilwa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portele

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scail

A:Authors: Schleich, S.; Schreier, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terplate, P.; Tognoni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipit, A.; Yamamoto, H.; Yamane, K.; Yasunaga, T.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danoch, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; MOID:98044033; PMID:9384377

A:Accession: F69592

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-157 <KUN>

A:Cross-references: GB:U99117; GB:AU09126; NID:q2634966; PIDN:CAH14613.1; PID:g26351

R:Belitsky, B.R.; Gustafsson, M.C.U.; Sorenshel, A.L.; von Wachenfeldt, C.

J. Bacteriol. 179, 5448-5457, 1997

A:Title: An 11p-11k gene of Bacillus subtilis involved in branched-chain amino acid

A:Reference number: Z22847; MOID:97411495; PMID:9287000

A:Accession: T44776

A>Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-157 <BEL>

A:Cross-references: EMBL:Y11043; NID:q1926275; PIDN:CAA71939.1; PID:q1926280

A:Experimental source: strain 1A1

C:Genetics:

A:Gene: azlB

C:Superfamily: regulatory protein asac

Query Match 71.4%; Score 35; DB 2; Length 157;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKRIGLSVSE 10
|:|||||: |
Db 27 LSKRIGLSVSE 36

RESULT 11
RGBCLR

leucine-responsive transcription regulator lrp [validated] - Escherichia coli (strain K12)

C:Species: Escherichia coli
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 01-Mar-2002
C:Accession: JH0412; 159240; H64827
R:Williams, D.A.; Ryan, C.W.; Platko, J.V.; Calvo, J.M.
J. Biol. Chem. 266, 10768-10774, 1991
A:Title: Characterization of lrp, an Escherichia coli regulatory protein that mediates a
A:Reference number: JH0412; MUID:91250365; PMID:2040596
A:Accession: JH0412
A:Molecule type: DNA
A:Residues: 1-164 <MIL>
A:Cross-references: GB:M5869; NID:g146664; PIDN:AAA24089.1; PID:g146665
R:Itto, K.; Kawakami, K.; Nakamura, Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 302-306, 1993
A:Title: Multiple control of Escherichia coli lysyl-tRNA synthetase expression involves
A:Reference number: 159240; MUID:9126364; PMID:7678344
A:Accession: 159240
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-164 <RES>
A:Cross-references: GB:D11105; NID:g216585; PIDN:BA01880.1; PID:g216586
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; et al.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64827
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-164 <BLAT>
A:Cross-references: GB:AE000191; GB:U000056; NID:g1787115; PIDN:AAC73975.1; PID:g1787116;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: lrp; alsB; llyR
C:Complex: homodimer [validated, MUID:91250369]
C:Function:
A:Description: activates a number of operons in response to the presence of exogenous leucine
C:Superfamily: regulatory protein asnc
C:Keywords: DNA binding; homodimer; transcription regulation
F:2-164/Product: 1 leucine-responsive regulatory protein #status experimental <MAY>
F:30-56/Region: helix-turn-helix motif

Query Match 71.4%; Score 35; DB 1; Length 164;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLS 8
DB 34 LSKRVGLS 41

RESULT 12
S59991
transcription regulator lrp, leucine-responsive - Serratia marcescens
C:Species: Serratia marcescens
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 26-Aug-1999
C:Accession: S59991
R:Calvo, J.M.
submitted to the EMBL Data Library, September 1993
A:Description: The amino acid sequence of lrp (leucine-responsive regulatory protein) in
A:Reference number: S59991
A:Accession: S59991
A:Molecule type: DNA
A:Residues: 1-164 <CAL>
A:Cross-references: EMBL:U02276; NID:g407920; PIDN:AA75466.1; PID:g407921
C:Genetics:
A:Gene: lrp
C:Function:
A:Description: activates a number of operons in response to the presence of exogenous leucine
C:Superfamily: regulatory protein asnc
C:Keywords: DNA binding; transcription regulation
F:30-56/Region: helix-turn-helix motif

Query Match 71.4%; Score 35; DB 2; Length 164;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLS 8
DB 34 LSKRVGLS 41

RESULT 13
S59993
transcription regulator lrp, leucine-responsive - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 26-Aug-1999
C:Accession: S59993
R:Calvo, J.M.
submitted to the EMBL Data Library, September 1993
A:Description: The amino acid sequence of lrp (leucine-responsive regulatory protein)
A:Reference number: S59991
A:Accession: S59993
A:Molecule type: DNA
A:Residues: 1-164 <CAL>
A:Cross-references: EMBL:U02273; NID:g407915; PIDN:AA75467.1; PID:g407916
C:Genetics:
A:Gene: lrp
C:Function:
A:Description: activates a number of operons in response to the presence of exogenous leucine
C:Superfamily: regulatory protein asnc
C:Keywords: DNA binding; transcription regulation
F:30-56/Region: helix-turn-helix motif

Query Match 71.4%; Score 35; DB 2; Length 164;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLS 8
DB 34 LSKRVGLS 41

RESULT 14
S59992
transcription regulator lrp, leucine-responsive - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 26-Aug-1999
C:Accession: S59992
R:Calvo, J.M.
submitted to the EMBL Data Library, September 1993
A:Description: The amino acid sequence of lrp (leucine-responsive regulatory protein)
A:Reference number: S59992
A:Accession: S59992
A:Molecule type: DNA
A:Residues: 1-164 <CAL>
A:Cross-references: EMBL:U02274; NID:g407917; PIDN:AA75465.1; PID:g407918
A:Note: The source is designated as Klebsiella ae-ogenes
C:Genetics:
A:Gene: lrp
C:Function:
A:Description: activates a number of operons in response to the presence of exogenous leucine
C:Superfamily: regulatory protein asnc
C:Keywords: DNA binding; transcription regulation
F:30-56/Region: helix-turn-helix motif

Query Match 71.4%; Score 35; DB 2; Length 164;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRIGLS 8
DB 34 LSKRVGLS 41

RESULT 15

B82142
leucine-responsive regulatory protein VC1904 [imported] - *Vibrio cholerae* (strain N16961)
C;Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence-revision 20-Aug-2000 #text-change 02-Feb-2001
C;Accession: B82142
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Yamthavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: B82142
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-164 <HR>
A;Cross-references: GB:AE004266; GB:AE003852; NID:99656435; PIDN:ANF95052.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1904
A;Map position: 1
C;Superfamily: regulatory protein asnc

Query Match 71.4%; Score 35; DB 2; Length 164;
Best Local Similarity 87.5%; Pred. NO. 15;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSKRTGLS 8
DB 34 LSKRVGLS 41

Search completed: January 15, 2003, 12:37:27
Job time : 16.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 Seconds

(without alignments)
72.584 Million cell updates/sec

Title: US-09-823-649A-6

Perfect score: 49

Sequence: 1 LSKRIGLSVSE 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	75.5	162	1 AROK_HELPY	P56073 helicobacte
2	36	73.5	548	1 TCPO_HUMAN	P50990 homo sapien
3	36	73.5	548	1 TCPO_MOUSE	P42932 mus musculu
4	36	73.5	914	1 SYA_PYPAR	Q8U425 pyrococcus
5	35	71.4	116	1 CHPB_ECOLI	P43647 escherichia
6	35	71.4	156	1 PYTR_AGRUU	Q44333 agrobacteri
7	35	71.4	157	1 AZLB_BACSU	007920 bacillus su
8	35	71.4	163	1 LRP_ECOLI	P19494 escherichia
9	35	71.4	163	1 LRP_KLEPN	P37424 klebsiella
10	35	71.4	163	1 LRP_SALTY	P37403 salmonella
11	35	71.4	163	1 LRP_SERMA	P37425 serratia ma
12	35	71.4	519	1 GALI_SCHPO	Q9Hdu2 schizosacch
13	35	71.4	585	1 RM04_SCHPO	Q9Hdu2 schizosacch
14	34	69.4	199	1 RS4_BACSU	P37505 bacillus su
15	34	69.4	201	1 YFAS_BACSU	P37505 bacillus su
16	34	69.4	374	1 CATB_PSEPU	P08310 pseudomonas
17	34	69.4	464	1 XKDK_BACSU	P54331 bacillus su
18	34	69.4	485	1 R16_MYXXA	P22072 myxococcus
19	34	69.4	664	1 ACES_ANOST	P56161 anopheles s
20	34	69.4	759	1 PPRU_THFAC	Q9H144 thermoplasma
21	34	69.4	795	1 CBBC_HUMAN	Q9H511 homo sapien
22	33	67.3	162	1 AROK_HELPY	Q9ZMS3 helicobacte
23	33	67.3	185	1 PTH_THFEN	Q8H757 thetmoanaer
24	33	67.3	336	1 Y625_METJA	Q58044 methanococc
25	33	67.3	337	1 YV76_METJA	Q58476 methanococc
26	33	67.3	379	1 CTRB1_ACTIVM	Q31946 acetoabact
27	33	67.3	507	1 Y62M_YEAST	P45295 saccharomyc
28	33	67.3	542	1 HMA2_CUTSA	P45295 cucumis sat
29	33	67.3	545	1 N05M_ALBCO	P46918 albinaria c
30	33	67.3	822	1 DEXT_STRSL	Q59979 streptococc
31	33	67.3	915	1 SYA_PYPAR	Q58033 pyrococcus
32	33	67.3	1034	1 CAPP_SYNY3	P74299 synechocyst
33	32	65.3	113	1 YNFB_ECOLI	P76170 escherichia

34	32	65.3	119	1 RNPA_LISIN	Q92644 listeria in
35	32	65.3	119	1 RNPA_LISMO	Q8Y311 listeria mo
36	32	65.3	138	1 RR9_ODOSI	P49497 odontella s
37	32	65.3	163	1 TPC_BRALA	P80322 branchiosto
38	32	65.3	166	1 LRP_HAEIN	P45265 haemophilus
39	32	65.3	188	1 YECW_ECOLI	P52007 escherichia
40	32	65.3	189	1 Y064_METJA	Q60376 methanococc
41	32	65.3	198	1 RS4_BACST	P81288 bacillus st
42	32	65.3	226	1 OCG1_AOUAE	Q66612 aquifex aco
43	32	65.3	231	1 VAIL_MYCPN	P75102 mycoplasma
44	32	65.3	300	1 SDB1_RAP	Q94192 rattus norv
45	32	65.3	315	1 FLIM_BUCAI	P57182 buchmera ap

ALIGNMENTS

RESULT 1	AROK_HELPY	STANDARD:	PRT:	162 AA.
ID	AROK_HELPY			
AC	P56073:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-0UN-2002 (Rel. 41, Last annotation update)			
DE	Shikimate kinase (EC 2.7.1.71) (SK).			
GN	AROK OR HP0157.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria: Proteobacteria: epsilon subdivision: Helicobacter group:			
OC	Helicobacter			
OX	NCBI_TaxID=210;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=26695 / ATCC 700392;			
RX	MEDLINE=97394467; PubMed=9252185;			
RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,			
RA	Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,			
RA	Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,			
RA	Lottus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,			
RA	McKenny K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,			
RA	Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,			
RA	Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,			
RA	Hayes W.S., Borodovsky M., Karpi P.D., Smith H.O., Fraser C.M.,			
RA	Venter J.C.;			
RT	"The complete genome sequence of the gastric pathogen Helicobacter			
RL	pylori.";			
RL	Nature 388:539-547(1997).			
CC	-1- CATALYTIC ACTIVITY: ATP + shikimate = ADP + shikimate-3-phosphate.			
CC	-1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;			
CC	-1- STEP.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (probable).			
CC	-1- SIMILARITY: BELONGS TO THE SHIKIMATE KINASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: AE000536; AAD07220.1; -			
DR	HSSP: P10880; ISHK.			
DR	TIGR: HP0157; -			
DR	InterPro: IPR000623; Shik_kinase.			
DR	Pfam: PF01202; SKI; 1.			
DR	PRINTS: PR01100; SHIKIMTKINASE.			
DR	PROSITE: PS01128; SHIKIMATE_KINASE; 1.			
KW	Aromatic amino acid biosynthesis; Transferase; Kinase; ATP-binding;			
KW	Complete proteome.			
FT	SE BIND			
FT	8			
FT	15			
FT	ATP (POTENTIAL).			
FT	SEQUENCE 162 AA; 18411 MW; 3D4B5F74E1C958FE CRC64;			
Query Match	75.5%;	Score 37;	DB 1;	Length 162;

Best Local Similarity 63.6%; Pred. No. 3.6;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISKRGISVSB 11
:1:111111
Db 36 ISEYGLSVSE 46

RESULT 2

TCPO_HUMAN STANDARD; PRT; 548 AA.
AC P50990;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE T-complex protein 1, theta subunit (TCP-1-theta) (CCT-theta).
GN CCT8 OR CCT9 OR KIA0002.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=bone marrow;
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawabayashi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT Prediction of the coding sequences of unidentified human genes. I.
RT The coding sequences of 40 new genes (KIA0001-KIA0040) deduced by
RT analysis of randomly sampled cDNA clones from human immature myeloid
RT cell line KG-1.
RL DNA Res. 1:27-35(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127533; PubMed=8590283;
RA Yamazaki M., Ono A., Watanabe K., Sasaki K., Tashiro H., Nomura T.;
RT Nucleotide sequence surrounding the locus marker D21S246 on human
RT chromosome 21.
RL DNA Res. 2:187-189(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Taudien S., Dagand E., Delabar J., Nordsiek G., Drescher B., Weber J.,
RA Schattvey R., Menzel U., Yaspo M.-L., Rosenthal A.;
RT Submitted (FEB-1999) to the EMBL/GenBank/DBD databases.
RL [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=2039799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Toki Y., Choi D.-K., Sueda E.,
RA Ohki M., Takagi T., Sasaki Y., Taudien S., Blechschmidt K., Folley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shitani A., Sasaki T., Nacamine K., Mitsuyama S., Antolacakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandl P.,
RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Blecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Lehmeyer S., Bozrym K., Gardiner K., Nictelic D., Francis F.,
RA Wehner H., Reinhardt R., Yaspo M.-L.;
RT The DNA sequence of human chromosome 21.
RL Nature 405:311-319(2000).
RN [5]
RP SEQUENCE OF 33-65 FROM N.A.
RX MEDLINE=95197008; PubMed=7890169;
RA Kubota H., Hynes G., Willison K.;
RT The eighth Cct gene, Cct9, encoding the theta subunit of the
RT cytosolic chaperonin containing TCP-1.
RL Gene 154:231-236(1995).
CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
CC ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC ACTIN AND TUBULIN.
CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 kDa THAT
CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D13627; BAA02792.1;
DR EMBL: D42052; BAA07652.1; AIN SEQ.
DR EMBL: AF129075; NOT ANNOTATED_CDS.
DR EMBL: A1163249; CAB90433.1;
DR EMBL: Z37163; CAB85520.1;
DR JSSP: P48424; 1A6D.
DR Genew: HGNC:1623; CCT8.
DR InterPro: IPR002194; Chaperonin_TCP-1.
DR InterPro: IPR001844; Chaperin_Cpn60.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; cpn60_TCP1.1.
DR PRINTS: PR00296; CHAPERONIN60.
DR PRINTS: PR00304; TCOMPLEXTCP1.
DR PROSITE: PS00750; TCP1_1; 1.
DR PROSITE: PS00751; TCP1_2; 1.
DR PROSITE: PS00995; TCP1_3; 1.
KW Chaperone; ATP-binding; Multigene family.
FT CONFLICT 50 N->K (IN REF. 1 AND 2).
FT CONFLICT 391 A->V (IN REF. 1).
SQ SEQUENCE 548 AA; 59620 MW; 566A622BC2D15E9 CRC64;
Query Match 73.5%; Score 36; DB 1; Length 548;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 RIGISVSE 11
:1:111111
Db 120 RIGISVSE 127
RESULT 3
TCPO_MOUSE STANDARD; PRT; 548 AA.
AC P42932;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 35, Last annotation update)
DE T-complex protein 1, theta subunit (TCP-1-theta) (CCT-theta).
GN CCT8 OR CCT9.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclungnathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95197008; PubMed=7890169;
RA Kubota H., Hynes G., Willison K.;
RT The eighth Cct gene, Cct9, encoding the theta subunit of the
RT cytosolic chaperonin containing TCP-1.
RL Gene 154:231-236(1995).
CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
CC ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC ACTIN AND TUBULIN.
CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 kDa THAT
CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: Z37164; CAA85521.1; -
 CC HSSP: P48424; 1A6D.
 DR MGD: MGI:107183; Cct8.
 DR InterPro: IPR002194; Chaperonin_TCP-1.
 DR InterPro: IPR002423; Cpn60/TCP-1.
 DR Pfam: PF00118; cpn60_TCP1.1.
 DR PRINTS: PR00304; TCOMPLXTCPL.
 DR PROSITE: PS00750; TCP1_1; 1.
 DR PROSITE: PS00751; TCP1_2; 1.
 DR PROSITE: PS00995; TCP1_3; 1.
 KW Chaperone; ATP-binding; Multigene family.
 SQ SEQUENCE 548 AA; 59555 MW; 4B5265250CFPIFE2 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIGLVSSE 11
 DB 120 RIGLVSSE 127

RESULT 4

SYA_PYRFU STANDARD; PRT; 914 AA.

AC 080425;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (Alans).
 GN ALAS OR PF0270.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -1 CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
 CC diposphate + L-alanyl-tRNA(Ala).
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1 SIMILARITY: BELONGS TO CLASS-TI AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE010151; AAL80394.1; -
 DR PROSITE: PS50860; AA-TRNA_LIGASE_ILALA.1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 914 AA; 104906 MW; 467F593731BD8BC1 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 914;
 Best Local Similarity 54.5%; Pred. No. 31;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSKRIGLVSSE 11
 DB 317 VAKRIGLVSSE 327

RESULT 5

CHPE_ECOLI STANDARD; PRT; 116 AA.

AC P33647;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Peak-like protein 2.
 GN CHPE OR CHPB OR H4225.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=94042847; PubMed=8226627;
 RA Masuda Y., Miyakawa K., Nishimura Y., Ohtsubo E.;
 RT "ChpA and chpB, Escherichia coli chromosomal homologs of the pem
 RT locus responsible for stable maintenance of plasmid R100."
 RL J. Bacteriol. 175:6850-6856(1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes."
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN [3]

RP SEQUENCE OF 1-30 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=86278038; PubMed=3525538;
 RA Weiss D.L., Johnson D.I., Welch H.L., Somerville R.L.;
 RT "Structural analysis of the *hcr* locus of Escherichia coli K12."
 RL J. Biol. Chem. 261:9966-9971(1986).
 RN [4]

RP SEQUENCE OF 20-116 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=89053923; PubMed=2848015;
 RA Lahli R., Pliskaeranta T., Valve E., Iila I., Kukko-Kalske E.,
 RA Helonen J.;
 RT "Cloning and characterization of the gene encoding inorganic
 RT pyrophosphatase of Escherichia coli K-12."
 RL J. Bacteriol. 170:5901-5907(1988).
 RN [5]

RP GENE MAPPING;
 RX MEDLINE=94364970; PubMed=808180;
 RA Masuda Y., Ohtsubo E.;
 RT "Mapping and disruption of the *chpA* locus in Escherichia coli."
 RL J. Bacteriol. 175:5861-5863(1994).
 CC -1 PUNCTION: MAY BE INVOLVED IN THE REGULATION OF CELL GROWTH, IT ACT
 CC AS A GROWTH INHIBITOR. BOTH CHPS AND CHPB BIND TO THE PROMOTER
 CC REGION OF THE CHPB OPERON TO AUTOREGULATE THEIR SYNTHESIS.
 CC -1 SIMILARITY: BELONGS TO THE PEAK FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D16451; BAA03920.1; -
 DR EMBL: M14018; -; NOT_ANNOTATED_CDS.
 DR EMBL: U14003; AAA97122.1; -
 DR EMBL: AE000494; AAC7182.1; -
 DR EMBL: M23550; -; NOT_ANNOTATED_CDS.
 DR PIR: D49339; D49339.
 DR EcGene: EG12096; ChpB.
 DR InterPro: IPR003477; PeakK.

DR Pfam: PF02452; PenK; 1.
 KW DNA-binding: Complete proteome.
 SQ SEQUENCE 116 AA; 12492 MW; 4D61F3E3029CF58 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 116;
 Best Local Similarity 63.6%; Pred. No. 6.3;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSKRIGLSVSE 11
 |:|||||:
 DB 92 LAKRIGLADE 102

RESULT 6
 PTRR_AGRPU STANDARD; PRT; 156 AA.
 ID PTRR_AGRPU
 AC Q4433;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 GN Proline dehydrogenase transcriptional activator.
 OS Agrobacterium tumefaciens.
 OG Plasmid pAK10.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 OX NCBI_Taxid=358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R10;
 RX MEDLINE=96178949; PubMed=8606160;
 RA Cho K., Fugua C., Martin B.S., Minns S.C.;
 RT "Identification of Agrobacterium tumefaciens genes that direct the
 complete catabolism of octopine.";
 RL J. Bacteriol. 178:1872-1880(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE PUTA GENE IN RESPONSE
 TO PROLINE.
 CC -1- SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
 REGULATORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U39263; AAC3979.1;
 DR InterPro: IPR000485; ASNC_trans_reg.
 DR Pfam: PF01037; ASNC_trans_reg; 1.
 DR PRINTS: PR00033; HTASNC.
 DR SMART: SM00344; HTH_ASNC; 1.
 DR PROSITE: PS00519; HTH_ASNC_FAMILY; 1.
 KW Transcription regulation; DNA-binding; Activator; Plasmid.
 FT DNA_BIND 29 48 H-R-H MOTIF (POTENTIAL).
 SQ SEQUENCE 156 AA; 17611 MW; PFC04E5B594C3881 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 156;
 Best Local Similarity 87.5%; Pred. No. 8.4;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSKRIGLS 8
 |:|||||:
 DB 32 LSKRIGLS 39

RESULT 7
 AZLB_BACSU STANDARD; PRT; 157 AA.
 ID AZLB_BACSU
 AC Q07920;
 DT 15-JUN-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcriptional regulator azlB.
 CN AZLB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / HSC1A1;
 RX MEDLINE=97431495; PubMed=9287000;
 RA Belitsky B.R., Gustafsson M.C.U., Sonenshein A.L., von Wachenfeldt C.;
 RT "An lrp-like gene of Bacillus subtilis involved in branched-chain
 amino acid transport.";
 RL J. Bacteriol. 179:5448-5457(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97453479; PubMed=9308178;
 RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
 R. Dusterhoelt A., Ehrlich S.D.;
 RT "Sequence of the Bacillus subtilis genome region in the vicinity of
 the lev operon reveals two new extracytoplasmic function RNA
 polymerase sigma factors Sigv and Sigz.";
 RL Microbiology 143:2939-2943(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S.,
 Borriani R., Borrer L., Brans A., Braun M., Briggell S.C., Bron S.,
 Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Conerton J.F., Cummings N.J., Daniel R.A.,
 Denizot F., Devine K.M., Dusterhoelt A., Ehrlich S.D., Emmerson P.T.,
 Entian K.D., Farrington J., Fabret C., Ferrari E., Foulger D.,
 Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
 Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
 Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Kieffer-Hlanchard M., Klein C.,
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 Kunita K., Lapilus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 Priesner K., Puje P., Purnelle B., Kapoport G., Key M., Keymolds S.,
 Rieger M., Rivolta C., Kocha E., Kocher B., Kose M., Sadale Y.,
 Sato T., Stanlan E., Schleich S., Schroeter R., Scott F.,
 Sekiguchi J., Sekowska A., Siroi S.J., Serror P., Shin B.S., Soldo B.,
 Sorokin A., Tarcioni R., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassart A.,
 Vialti A., Wambolt R., Wedler E., Wedler H., Wellenzeger T.,
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF THE AZLBCD OPERON INVOLVED
 IN BRANCHED-CHAIN AMINO ACID TRANSPORT.
 CC -1- SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
 REGULATORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: Y11043; CAA71939.1; -

RN [5]

DR EMBL; D90726; BA

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is and for commercial use. The content of this entry is not removed, usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: D90727; BAA5621.1; -
 DR EMBL: AE005278; AAG55376.1; -
 DR EMBL: AP002553; BAB34397.1; -
 DR EMBL: U02272; AAG75439.1; -
 DR PIR: JH0412; RGECLR.
 DR ECO2DBASE: 1015.1; 6TH EDITION.
 DR Ecogene: EG10547; Lrp.
 DR InterPro: IPR000485; ASNC.trans.reg.
 DR Pfam: PF01037; ASNC.trans.reg; 1.
 DR PRINTS: PR00033; HTASNC.
 DR SMART: SM00344; HTH_ASNF; 1.
 DR PROSITE: PS00519; HTH_ASNF_FAMILY; 1.
 KW DNA-binding; Transcription regulation; Activator; Complete proteome.
 FT INIT MET 0
 FT DNA_BIND 30 49 H-T-H MOTIF (POTENTIAL).
 FT VARIANT 113 113 D -> E (IN LRP-1 MUTANT).
 SQ SEQUENCE 163 AA; 18755 MW; B008349588F7B255 CRC64;
 Query Match 71.4%; Score 35; DB 1; Length 163;
 Best Local Similarity 87.5%; Pred. No. 8.8;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSRIGLS 8
 |||||
 Db 33 LSKRWGLS 40

RESULT 9
 LRP_KLEPN STANDARD; PRT: 163 AA.
 AC P37424;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Leucine-responsive regulatory protein.
 GN Lrp.
 OS Klebsiella pneumoniae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95189743; PubMed-7883720;
 RA Friedberg D., Plakko J.V., Tyler B., Calvo J.M.;
 RT "The amino acid sequence of Lrp is highly conserved in four enteric
 microorganisms.";
 RL J. Bacteriol. 177:1624-1626(1995).
 CC -!- FUNCTION: MEDIATES A GLOBAL RESPONSE TO LECICINE. EXOGENOUS LECICINE
 AFFECTS THE EXPRESSION OF A NUMBER OF DIFFERENT OPERONS. LRP
 MEDIATES THIS EFFECT FOR AT LEAST SOME OF THESE OPERONS. FOR
 EXAMPLE IT IS REGULATOR OF THE BRANCHED-CHAIN AMINO ACID TRANSPORT
 GENES.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
 REGULATORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: U02274; AAG75465.1; -
 DR InterPro: IPR000485; ASNC.trans.reg.
 DR Pfam: PF01037; ASNC.trans.reg; 1.
 DR PRINTS: PR00033; HTASNC.
 DR SMART: SM00344; HTH_ASNF; 1.
 DR PROSITE: PS00519; HTH_ASNF_FAMILY; 1.
 KW DNA-binding; Transcription regulation; Activator.
 FT INIT MET 0
 FT DNA_BIND 30 49 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 163 AA; 18741 MW; B6D5EP95B8F3A645 CRC64;

SQ SEQUENCE 163 AA; 18741 MW; B6D5EP95B8F3A645 CRC64;
 Query Match 71.4%; Score 35; DB 1; Length 163;
 Best Local Similarity 87.5%; Pred. No. 8.8;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSRIGLS 8
 |||||
 Db 33 LSKRWGLS 40

RESULT 10
 LRP_SALTY STANDARD; PRT: 163 AA.
 ID LRP_SALTY
 AC P37403; 087635;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-responsive regulatory protein.
 GN LRP OR STM0959.
 OS Salmonella typhimurium, and
 OS Klebsiella aerogenes.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602, 28451;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S. typhimurium;
 RX MEDLINE-95189743; PubMed-7883720;
 RA Friedberg D., Plakko J.V., Tyler B., Calvo J.M.;
 RT "The amino acid sequence of Lrp is highly conserved in four enteric
 microorganisms.";
 RL J. Bacteriol. 177:1624-1626(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S. typhimurium; STRAIN-LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Muthaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-K. aerogenes; STRAIN-W70;
 RX MEDLINE-99121072; PubMed-9922277;
 RA Jones B.K., Bender R.A.;
 RT "Two roles for the leucine-responsive regulatory protein in expression
 of the alanine catabolic operon (dadA) in Klebsiella aerogenes.";
 RL J. Bacteriol. 181:1054-1058(1999).
 CC -!- FUNCTION: MEDIATES A GLOBAL RESPONSE TO LECICINE. EXOGENOUS LECICINE
 AFFECTS THE EXPRESSION OF A NUMBER OF DIFFERENT OPERONS. LRP
 MEDIATES THIS EFFECT FOR AT LEAST SOME OF THESE OPERONS. FOR
 EXAMPLE IT IS REGULATOR OF THE BRANCHED-CHAIN AMINO ACID TRANSPORT
 GENES.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE ASNC FAMILY OF TRANSCRIPTIONAL
 REGULATORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: U02273; AAG75467.1; -
 DR EMBL: AE008741; AAL19894.1; -
 DR EMBL: AF090144; AAD12584.1; -

DR InterPro: IPR000241; RNA_methylase.
 DR InterPro: IPR000051; SAM_bind.
 DR InterPro: IPR004114; THUMP_dom.
 DR Pfam: PF01170; UPF020; 1.
 DR Pfam: PF02926; THUMP; 1.
 DR TIGRfams: TIGR01177; TIGR01177; 1.
 DR PROSITE: PS01261; UPF020; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 320 AA; 36673 MW; A879F320A8CD637 CRC64;

Query Match 67.3%; Score 33; DB 1; Length 320;
 Best Local Similarity 70.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Caps 0;

OY 2 AONNIXRKE 11
 |||||
 Db 9 ADNLEIARKE 18

RESULT 15

YGJ3_YEAST STANDARD; PRT; 917 AA.
 AC P53148; 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 104.8 kDa protein in PAN2-NUP145 intergenic region.
 GN YGI093W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97435481; PubMed=9290212;
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
 RT "Sequence analysis of 203 kilobases from saccharomyces cerevisiae
 chromosome VII.";
 RL Yeast 13:1077-1090(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z72615; CAA96799.1; -;
 DR SCD; S0003061; SPC105.
 KW Hypothetical protein.
 SQ SEQUENCE 917 AA; 104825 MW; 8B7ED1522916A319 CRC64;

Query Match 67.3%; Score 33; DB 1; Length 917;
 Best Local Similarity 54.5%; Pred. No. 40;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Caps 0;

OY 1 LAONNIXRKE 11
 |||||
 Db 594 LAENNTIKRE 604

Search completed: January 15, 2003, 11:21:36
 Job time : 9.28572 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 11:21:44 ; Search time 130.143 Seconds
(without alignments)
190.345 Million cell updates/sec

Title: US-09-823-649A-6
Sequence: 49
1 LSKRIGLSVSE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -f -
-O=/gen2.1/USPRO.spool/US09823649/runat.14012003.151001.29119/app_query.fasta.1.1393
-DB=N.Geneseq.101002 -OPMT=faststep -SUFFIX=ring -MINMATCH=0.1 -LOPCP=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.@CN.1.1.0.@runat.14012003.151001.29119 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOC -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	49	100.0	1830	AAQ24334	Mutant thermostabl
2	49	100.0	2073	AAQ24333	Mutant thermostabl
3	49	100.0	2265	AAQ24332	Mutant thermostabl
4	49	100.0	2403	AAQ24331	Mutant thermostabl
5	49	100.0	2568	AAQ28937	Encodes Taf DNA po
6	49	100.0	2571	AAQ24330	Mutant thermostabl
7	49	100.0	2679	AAQ24329	Mutant thermostabl
8	49	100.0	2679	AAQ28936	Encodes Asp37 Taf
9	49	100.0	4286	AAQ3917	Taf DNA polymerase
10	42	85.7	580073	AAQ58840	Mycoplasma genital
11	41	83.7	163319	AAQ22306	Arabidopsis thalia
12	39	79.6	471	ABA51767	Human foetal liver
13	39	79.6	471	ABA21536	Probe #62 for gene
14	39	79.6	471	AAK00075	Human brain expres
15	39	79.6	471	AAK25512	Human bone marrow
16	39	79.6	471	AAI10135	Probe #68 for gene
17	39	79.6	471	AAI31384	Probe #70 used to
18	39	79.6	471	AAI00076	Probe #67 used to
19	39	79.6	471	ABSO0080	Human genome deriv
20	39	79.6	668	ABA46582	Human breast cell
21	39	79.6	668	ABA64445	Human foetal liver
22	39	79.6	668	ABA31582	Probe #10048 for g
23	39	79.6	668	AAK12903	Human brain expres
24	39	79.6	668	AAK38630	Human bone marrow
25	39	79.6	668	AAI19430	Probe #9363 for ge
26	39	79.6	668	AAI44621	Probe #13167 used
27	39	79.6	668	AAI05155	Probe #1316 used t
28	39	79.6	668	ABSI2699	Human genome deriv
29	39	79.6	468	AAH52475	S. epidermidis ope
30	38	77.6	495	ABN92916	Staphylococcus epi
31	38	77.6	2122	AAH86153	DNA encoding novel
32	38	77.6	3472	AAH34107	S. epidermidis gen
33	38	77.6	4105	AAH54115	S. epidermidis gen
34	37	75.5	1314	ABK74034	Bacillus lichenifo
35	36	73.5	430	AAV86138	EST clone H306. H
36	36	73.5	460	AAZ42996	Human 5' EST isola
37	36	73.5	657	AAH56454	DNA encoding novel
38	36	73.5	1024	AAH7621	Human cancer assoc
39	36	73.5	1602	AAV65278	DNA sequence of S.
40	36	73.5	1785	AAV65288	DNA encoding a S.
41	36	73.5	1794	AAH86151	DNA encoding novel
42	36	73.5	1871	AAH39656	Renal cancer assoc
43	36	73.5	1843	ABK36206	cDNA sequence #397
44	36	73.5	1853	AAQ97801	Clone pthrela cod
45	36	73.5	1955	AAZ00464	Human secreted pro

ALIGNMENTS

RESULT 1
AAQ24334 standard; DNA: 1830 BP.

XX AAQ24334;

DT 22-OCT-1992 (first entry)

DE Mutant thermostable DNA polymerase PTAfi285.

XX 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR; ss.

XX Thermophilic africanus.

OS Key location/Qualifiers

FT Old_sequence 1..2

FT /*tag= a

FT /note= "nucleotides 4-852 deleted from the native

sequence."

FT M09206200-A.
 XX 16-APR-1992.
 PD 30-SEP-1991; 91WO-US07035.
 XX 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX (CETU) CETUS CORP.
 PA Abramson RD, Gelfand DH;
 PI WPI; 1992-150885/18.
 DR P-PSDB; AAR23174.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 CC The sequence is that of Thermosipho africanus polymerase DNA which
 CC has been mutated. The mutation designated pTAF1285 causes the
 CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. Thermostable DNA
 CC polymerases are useful in many recombinant DNA techniques, esp.
 CC nucleic acid amplification by PCR, self-sustained sequence
 CC replication (SSR) and high temp. DNA sequencing. The absence of
 CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
 XX
 SQ Sequence 1830 BP; 719 A; 206 C; 352 G; 553 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.897 Length: 1830
 Score: 49.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-823-649a-6 (1-11) x AAQ24334 (1-1830)
 QY 1 LeuSerTysArgIleGlyLeuSerValSerGlu 11
 DB 1369 CTTTCAAGAGGAAATGCTTAGTTCAGAG 1401
 RESULT 2
 AAQ24333
 ID AAQ24333 standard; DNA; 2073 BP.
 XX
 AC AAQ24333;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase pTAFd2-203.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
 XX
 OS Thermosipho africanus.
 XX
 FH Key Location/Qualifiers

old_sequence 1..2
 FT /tag= a
 FT /note= "nucleotides 4-603 deleted from the native
 FT sequence."
 PN M09206200-A.
 XX 16-APR-1992.
 PD 30-SEP-1991; 91WO-US07035.
 XX 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX (CETU) CETUS CORP.
 PA Abramson RD, Gelfand DH;
 PI WPI; 1992-150885/18.
 DR P-PSDB; AAR23173.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 CC The sequence is that of Thermosipho africanus polymerase DNA which
 CC has been mutated. The mutation designated pTAFd2-203 causes the
 CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. Thermostable DNA
 CC polymerases are useful in many recombinant DNA techniques, esp.
 CC nucleic acid amplification by PCR, self-sustained sequence
 CC replication (SSR) and high temp. DNA sequencing. The absence of
 CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
 XX
 SQ Sequence 2073 BP; 820 A; 226 C; 399 G; 628 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1.04 Length: 2073
 Score: 49.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-823-649a-6 (1-11) x AAQ24333 (1-2073)
 QY 1 LeuSerTysArgIleGlyLeuSerValSerGlu 11
 DB 1612 CTTTCAAGAGGAAATGCTTAGTTCAGAG 1644
 RESULT 3
 AAQ24332
 ID AAQ24332 standard; DNA; 2265 BP.
 XX
 AC AAQ24332;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase pTAF11.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
 XX

OS Thermosipho africanus.
XX
XX Key Location/Qualifiers
FH 1..2
FT old_sequence /tag= a
FT /note= "nucleotides 4-417 deleted from the native
sequence."
XX
XX MO9206200-A.
XX
XX 16-APR-1992.
XX
XX 30-SEP-1991; 91WO-US07035.
XX
XX 28-SEP-1990; 90US-0590213.
XX 28-SEP-1990; 90US-0590466.
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU) CETUS CORP.
XX
XX Abramson RD, Gelfand DH;
XX
XX WPI: 1992-150885/18.
XX P-PSDB; AAR23172.
XX
XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
XX Claim 11; Page 59; 185pp; English.
XX
XX The sequence is that of Thermosipho africanus polymerase DNA which
CC has been mutated. The mutation designated PTAf09 causes the
CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
CC exonuclease activity than the native enzyme. Thermostable DNA
CC polymerases are useful in many recombinant DNA techniques, esp.
CC nucleic acid amplification by PCR, self-sustained sequence
CC replication (SSR) and high temp. DNA sequencing. The absence of
CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
CC discrimination in a combined polymerase ligase chain reaction (PLCR)
CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
CC in enzymes used in homogeneous assays for the amplification and
CC detection of a target nucleic acid sequence. Mutation of the DNA
CC encoding particular regions of the enzymes can be used to prepare
CC a range of recombinant proteins having 5'-3' exonuclease activity
CC to a complete lack of activity.
CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
XX Sequence 2265 BP; 892 A; 245 C; 440 G; 688 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.16 Length: 2265
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-823-649A-6 (1-11) x AAQ24331 (1-2265)
QY 1 LeuserLysArgIleGlyLeuserValSerGlu 11
DB 1804 CTTTCAAGAGAAATGCTTAGTTCAGAG 1836
RESULT 4
ID AAQ24331 standard; DNA; 2403 BP.
XX
XX AAQ24331;
XX
XX 22-OCT-1992 (first entry)
XX
XX Mutant thermostable DNA polymerase PTAf09.

XX
XX 5'-3' exonuclease; PCR; amplification; SSR; sequencing; PCR; ss.
XX
XX Thermosipho africanus.
XX
XX Key Location/Qualifiers
FH 1..2
FT old_sequence /tag= a
FT /note= "nucleotides 4-279 deleted from the native
sequence."
XX
XX MO9206200-A.
XX
XX 16-APR-1992.
XX
XX 30-SEP-1991; 91WO-US07035.
XX
XX 28-SEP-1990; 90US-0590213.
XX 28-SEP-1990; 90US-0590466.
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU) CETUS CORP.
XX
XX Abramson RD, Gelfand DH;
XX
XX WPI: 1992-150885/18.
XX P-PSDB; AAR23171.
XX
XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
XX Claim 11; Page 59; 185pp; English.
XX
XX The sequence is that of Thermosipho africanus polymerase DNA which
CC has been mutated. The mutation designated PTAf09 causes the
CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
CC exonuclease activity than the native enzyme. Thermostable DNA
CC polymerases are useful in many recombinant DNA techniques, esp.
CC nucleic acid amplification by PCR, self-sustained sequence
CC replication (SSR) and high temp. DNA sequencing. The absence of
CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
CC discrimination in a combined polymerase ligase chain reaction (PLCR)
CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
CC in enzymes used in homogeneous assays for the amplification and
CC detection of a target nucleic acid sequence. Mutation of the DNA
CC encoding particular regions of the enzymes can be used to prepare
CC a range of recombinant proteins having 5'-3' exonuclease activity
CC to a complete lack of activity.
CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
XX Sequence 2403 BP; 947 A; 291 C; 461 G; 732 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.24 Length: 2403
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-823-649A-6 (1-11) x AAQ24331 (1-2403)
QY 1 LeuserLysArgIleGlyLeuserValSerGlu 11
DB 1942 CTTTCAAGAGAAATGCTTAGTTCAGAG 1974
RESULT 5
ID AAQ28937 standard; DNA; 2568 BP.
XX
XX AAQ28937;
XX
XX

DN 27-OCT-1992 (first entry)
XX
DE Encodes Taf DNA polymerase I lacking 5'-3' exonuclease activity.
XX
KW Thermostability; PCR; polymerase chain reaction;
XX thermophilic bacteria; Taf Pol I; deletion mutant; ss.
OS Thermosipho africanus.
XX
PN W09206202-A.
XX
PD 16-APR-1992.
XX
PF 26-SEP-1991; 91WO-US07076.
XX
PR 28-SEP-1990; 90US-0590490.
XX
PA (CETU) CETUS CORP.
XX
PI Abramson RD, Gelfand DH, Greenfield L, Lawyer FC, Reichert FU,
XX
DR WPI; 1992-150867/18.
DR P-PSDB; AAR23122.
XX
PT Thermostable DNA polymerase from Thermosipho africanus - prepd.
PT by purification, from cells or by expression of Taf polymerase gene
PT in host cells
XX
PS Claim 21; Page 70; 80pp; English.
XX
CC The sequence coding for a thermostable DNA polymerase was isolated
CC from chromosomal DNA of Thermosipho africanus (Taf). The polymerase
CC (see AAR23122) has 5'-3' exonuclease activity. Deletion of codon 1-37
CC results in a DNA polymerase which lacks the 5'-3' exonuclease
CC activity. See AAQ23917 for the wild-type Taf Pol I gene and AAQ28936 for
CC another preferred mutant.
XX
SQ Sequence 2568 BP; 1012 A; 278 C; 494 G; 784 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.35 Length: 2568
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-823-649a-6 (1-11) x AAQ28937 (1-2568)
QY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db 2107 CTTTCAAGAGAAATGGCTCTAGTCTTCACAG 2139
RESULT 6
AAQ24330
ID AAQ24330 standard; DNA; 2571 BP.
XX
AC AAQ24330;
XX
DN 22-OCT-1992 (first entry)
XX
DE Mutant thermostable DNA polymerase pTAFd2-37.
XX
KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR; ss.
XX
OS Thermosipho africanus.
XX
PH Key Location/Qualifiers
FT 1..2
FT old_sequence /tag= a
FT /note= "nucleotides 4-111 deleted from the native
FT sequence."

PN W09206200-A.
XX
PD 16-APR-1992.
XX
PF 30-SEP-1991; 91WO-US07035.
XX
PR 28-SEP-1990; 90US-0590213.
XX
PR 28-SEP-1990; 90US-0590466.
XX
PR 28-SEP-1990; 90US-0590490.
XX
PA (CETU) CETUS CORP.
XX
PI Abramson RD, Gelfand DH;
XX
DR WPI; 1992-150865/18.
DR P-PSDB; AAR23170.
XX
PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT activity - having conserved regions mutated or deleted, for use
PT in e.g. PCR, sequencing and detection assays
XX
PS Claim 11; Page 59; 185pp; English.
XX
CC The sequence is that of Thermosipho africanus polymerase DNA which
CC has been mutated. The mutation designated pTAFd2-37 causes the
CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
CC exonuclease activity than the native enzyme. Thermostable DNA
CC polymerases are useful in many recombinant DNA techniques, esp.
CC nucleic acid amplification by PCR, self-sustained sequence
CC replication (SSR) and high temp. DNA sequencing. The absence of
CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
CC discrimination in a combined polymerase ligase chain reaction (PCR)
CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
CC in enzymes used in homogeneous assays for the amplification and
CC detection of a target nucleic acid sequence. Mutation of the DNA
CC encoding particular regions of the enzymes can be used to prepare
CC a range of recombinant proteins having 5'-3' exonuclease activity
CC to a complete lack of activity.
XX
SQ Sequence 2571 BP; 1014 A; 278 C; 494 G; 785 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.35 Length: 2571
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-823-649a-6 (1-11) x AAQ24330 (1-2571)
QY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db 2110 CTTTCAAGAGAAATGGCTCTAGTCTTCACAG 2142
RESULT 7
AAQ24329
ID AAQ24329 standard; DNA; 2679 BP.
XX
AC AAQ24329;
XX
DN 22-OCT-1992 (first entry)
XX
DE Mutant thermostable DNA polymerase from Thermosipho africanus.
XX
KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR; ss.
XX
OS Thermus thermophilus.
XX
PH Key Location/Qualifiers
FT 110
FT old_sequence /tag= a

```

FT      old_sequence /note= "G in native sequence"
FT      111
FT      /*tag= b
FT      /note= "A in native sequence"
XX
PN      W09206200-A.
XX
PD      16-APR-1992.
XX
PF      30-SEP-1991; 91WO-US07035.
XX
PR      28-SEP-1990; 90US-0590213.
PR      28-SEP-1990; 90US-0590466.
PR      28-SEP-1990; 90US-0590490.
XX
PA      (CETU ) CETUS CORP.
XX
PI      Abramson RD, Gelfand DH;
XX
DR      WP1; 1992-150885/18.
DR      P-PSDB; AAR23169.
XX
PT      Thermostable DNA polymerases with altered 5'-3' exo nuclease
PT      activity - having conserved regions mutated or deleted, for use
PT      in e.g. PCR, sequencing and detection assays
XX
PS      Claim 11; Page 59; 185pp; English.
XX
CC      The sequence is that of Thermosipho africanus polymerase DNA which has
CC      been mutated. The mutation causes the polymerase enzyme produced
CC      to exhibit a different amt. of 5'-3' exonuclease activity than the
CC      native enzyme. Thermostable DNA polymerases are useful in many
CC      recombinant DNA techniques, esp. nucleic acid amplification by
CC      PCR, self-sustained sequence replication (SSR) and high temp. DNA
CC      sequencing. The absence of 5'-3' nuclease activity may facilitate
CC      higher sensitivity allelic discrimination in a combined polymerase
CC      11gase chain reaction (PLCR) assay. An enhanced amt. of 5'-3'
CC      exonuclease activity may be desirable in enzymes used in homogeneous
CC      assays for the amplification and detection of a target nucleic acid
CC      sequence. Mutation of the DNA encoding particular regions of the
CC      enzymes can be used to prepare a range of recombinant proteins having
CC      5'-3' exonuclease activity to a complete lack of activity.
CC      See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
SQ      Sequence 2679 BP; 1045 A; 295 C; 515 G; 824 T; 0 other;

Alignment Scores:
Pred. No.: 1.42 Length: 2679
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-823-649a-6 (1-11) x AAQ24329 (1-2679)

QY      1 LeuSerIysaRgIleGlyLeuSerValSerGlu 11
DB      2218 CTTCAAGAGAGATTGCTTAGTCTTCAGAG 2250

RESULT 8
AAQ28936
ID      AAQ28936 standard; DNA: 2679 BP.
AC      AAQ28936;
XX
XX      27-OCT-1992 (first entry)
DE      Encodes Asp37 Taf DNA polymerase I lacking 5'-3' exonuclease activity.
XX
XX      Thermostability; PCR; polymerase chain reaction;
XX      thermophilic bacteria; Taf Pol I; mutant; ss.
XX

```

```

OS      Thermosipho africanus.
XX
XX      Key Location/Qualifiers
FT      mutation 109..111
FT      /*tag= a
FT      /note= "Gly codon changed to Asp"
XX
PN      W09206202-A.
XX
PD      16-APR-1992.
XX
PF      26-SEP-1991; 91WO-US07076.
XX
PR      28-SEP-1990; 90US-0590490.
XX
PA      (CETU ) CETUS CORP.
XX
PI      Abramson RD, Gelfand DH, Greenfield L, Lawyer FC, Reichert FI;
XX
DR      WP1; 1992-150887/18.
DR      P-PSDB; AAR23122.
XX
PT      Thermostable DNA polymerase from Thermosipho africanus - prepd.
PT      by purification from cells or by expression of Taf polymerase gene
PT      in host cells
XX
PS      Claim 20; Page 70; 80pp; English.
XX
CC      The sequence coding for a thermostable DNA polymerase was isolated
CC      from chromosomal DNA of Thermosipho africanus (Taf). The polymerase
CC      (see AAR23122) has 5'-3' exonuclease activity. Mutation of the codon
CC      specifying Gly at position 37 (i.e. GGA) to an Asp codon results in
CC      a DNA polymerase which lacks the 5'-3' exonuclease activity.
CC      See AAQ23917 for the wild-type Taf Pol I gene and AAQ28937 for another
CC      preferred mutant.
XX
SQ      Sequence 2679 BP; 1045 A; 295 C; 515 G; 823 T; 1 other;

Alignment Scores:
Pred. No.: 1.42 Length: 2679
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-823-649a-6 (1-11) x AAQ28936 (1-2679)

QY      1 LeuSerIysaRgIleGlyLeuSerValSerGlu 11
DB      2218 CTTCAAGAGAGAAATGCTTAGTCTTCAGAG 2250

RESULT 9
AAQ23917
ID      AAQ23917 standard; DNA: 4286 BP.
AC      AAQ23917;
XX
XX      27-OCT-1992 (first entry)
DE      Taf DNA polymerase I coding sequence.
XX
XX      Thermostability; PCR; polymerase chain reaction;
XX      thermophilic bacteria; Taf Pol I; ss.
XX
OS      Thermosipho africanus.
XX
XX      Key Location/Qualifiers
FT      CDS 298..2976
FT      /*tag= a
FT      /product= Polymerase_I
XX
PN      W09206202-A.

```

```

XX 16-APR-1992.
XX
XX 26-SEP-1991; 91WO-US07076.
XX
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU ) CETUS CORP.
XX
XX Abramson RD, Gelfand DH, Greenfield L, Lawyer FC, Reichert FL,
XX WPI, 1992-150887/18.
XX P-PSDB; AAK23122.
XX
XX Thermostable DNA polymerase from Thermophilus africanus - prep'd.
XX by purification from cells or by expression of Taf polymerase gene
XX in host cells
XX
XX Claim 8; Page 6; 80pp; English.
XX
XX Chromosomal DNA from Thermophilus africanus (Taf) was PCR-amplified
XX with degenerate primers corresponding to the amino acid sequences
XX of conserved regions of known thermostable polymerases. When
XX specific PCR products of a similar size to the product generated
XX using Taf chromosomal DNA were produced, the PCR fragments were
XX cloned and sequenced. Fragments with sequences which encoded
XX regions of amino acid homology to known thermostable polymerases
XX were identified. The cloned PCR products were used as probes to
XX screen a genomic Southern blot. The full-length Taf coding sequence
XX was then compiled from various clones. See also AAQ23918-Q23961.
XX
XX SQ Sequence 4286 BP; 1623 A; 470 C; 847 G; 1346 T; 0 other;

Alignment Scores:
Pred. No.: 2 48 Length: 4286
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-823-649A-6 (1-11) x AAQ23917 (1-4286)
QY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db 2515 CTTCAAGAGAAATTGCTTACTGTTTCAGAG 2547
RESULT 10
AAT58840
ID AAT58840 standard; DNA: 580073 BP.
XX
XX AAT58840;
XX
XX 27-MAR-1997 (first entry)
XX
XX Mycoplasma genitalium genome.
XX
XX M. genitalium; DNA: DNA gyrase; origin of replication;
XX megabase shotgun sequencing method; open reading frame; ORF; ss.
XX
XX Mycoplasma genitalium.
XX
XX Key Location/Qualifiers
XX CDS 8552..9184
XX /tag= a
XX /label= MG006
XX /note= "Previously identified as MORF-20076, the
XX encoded protein shows 27.59 percentage
XX identity to thymidylate kinase (CDC8)
XX from Saccharomyces cerevisiae"
XX
XX CDS 11252..12040
XX /tag= b
XX /label= MG009

```

```

FT /note= "Previously identified as MORF-20078, the
FT encoded protein shows 35.43 percentage
FT identity to the Bacillus subtilis hypothetical
FT protein covered in accession number
FT GB:D6185_102"
FT
FT CDS 12069..12725
FT /tag= c
FT /label= MG010
FT /note= "Previously identified as MORF-20079, the
FT encoded protein shows 25.73 percentage
FT identity to DNA primase (dnaE) from
FT Clostridium acetobutylicum"
FT
FT CDS complement (13570..14247)
FT /tag= d
FT /label= MG012
FT /note= "Previously identified as MORF-20080, the
FT encoded protein shows 31.50 percentage
FT identity to the ribosomal protein S6
FT modification protein (rimK) from Escherichia
FT coli"
FT
FT CDS complement (14396..15217)
FT /tag= e
FT /label= MG013
FT /note= "Previously identified as MORF-19823, MORF-20080
FT and MORF-20081, the encoded protein shows 33.04
FT percentage identity to 5,10-methylene-tetra-
FT hydrofolate dehydrogenase (fold) from E. coli"
FT
FT CDS 17474..19243
FT /tag= f
FT /label= MG015
FT /note= "Previously identified as MORF-20084, the
FT encoded protein shows 32.23 percentage
FT identity to transport ATP-binding protein
FT (msbA) from E. coli"
FT
FT CDS 26478..27344
FT /tag= g
FT /label= MG023
FT /note= "Previously identified as MORF-20092, the
FT encoded protein shows 45.96 percentage
FT identity to fructose-bisphosphate aldolase
FT (lsf) from B. subtilis"
FT
FT CDS 27345..28448
FT /tag= h
FT /label= MG024
FT /note= "Previously identified as MORF-19826 and
FT MORF-20093, the encoded protein shows 46.84
FT percentage identity to GTP-binding protein
FT from E. coli"
FT
FT CDS 36987..38978
FT /tag= i
FT /label= MG032
FT /note= "Previously identified as MORF-20099, the
FT encoded protein shows 26.82 percentage
FT identity to ATP-dependent nuclease (addA)
FT from B. subtilis"
FT
FT CDS 39242..39904
FT /tag= j
FT /label= MG033
FT /note= "Previously identified as MORF-20100, the
FT encoded protein shows 35.90 percentage
FT identity to glycerol uptake facilitator
FT (glpF) from B. subtilis"
FT
FT CDS complement (39873..40514)
FT /tag= k
FT /label= MG034
FT /note= "Previously identified as MORF-20101, the
FT encoded protein shows 48.13 percentage
FT identity to thymidylate kinase (tdk)
FT from B. subtilis"
FT
FT CDS 40543..41787
FT /tag= l
FT /label= MG035
FT /note= "Previously identified as MORF-20102, the

```

FT encoded protein shows 30.71 percentage
FT identity to histidyl-tRNA synthetase (hiss)
FT from *Mycobacterium leprae*"
CDS complement (44751..46277)
FT /tag= m
FT /label= MG038
FT /note= "Previously identified as MORF-20105, the
FT encoded protein shows 46.83 percentage
FT identity to glycerol kinase (glpk)
FT from *E. coli*"
CDS complement (46268..47422)
FT /tag= n
FT /label= MG039
FT /note= "Previously identified as MORF-19831 and
FT MORF-20106, the encoded protein shows 43.20
FT percentage identity to glycerol-3-phosphate
FT dehydrogenase (GUT2) from *S. cerevisiae*"
FT 49377..49643
FT /tag= o
FT /label= MG041
FT /note= "The encoded protein shows 48.86 percentage
FT identity to phosphohistidinoprotein-hexose
FT phosphotransferase (ptsH) from *Mycoplasma*
FT *capricolum*"
CDS 50060..51520
FT /tag= p
FT /label= MG042
FT /note= "Previously identified as MORF-19832 and
FT MORF-20108, the encoded protein shows 41.92
FT percentage identity to spermidine/
FT putrescine transport *Arp*-binding protein
FT (pota) from *E. coli*"
CDS 51525..52382
FT /tag= q
FT /label= MG043
FT /note= "Previously identified as MORF-20110, the
FT encoded protein shows 26.51 percentage
FT identity to spermidine/putrescine transport
FT system permease protein (potB) from *E. coli*"
CDS 52366..53220
FT /tag= r
FT /label= MG044
FT /note= "Previously identified as MORF-20111, the
FT encoded protein shows 29.45 percentage
FT identity to spermidine/putrescine transport
FT system permease protein C (potC) from *E. coli*"
CDS 54658..55605
FT /tag= s
FT /label= MG046
FT /note= "Previously identified as MORF-20112, the
FT encoded protein shows 36.60 percentage
FT identity to sialoglycoproteinase (gcp)
FT from *Pasteurella haemolytica*"
CDS complement (56970..58310)
FT /tag= t
FT /label= MG048
FT /note= "Previously identified as MORF-19834,
FT MORF-20114 and MORF-20115, the encoded protein
FT shows 43.02 percentage identity to signal
FT recognition particle protein (frh) from *B.*
FT *subtilis*"
CDS 58117..59079
FT /tag= u
FT /label= MG049
FT /note= "Previously identified as MORF-20114 and
FT MORF-20115, the encoded protein shows 44.78
FT percentage identity to purine-nucleoside
FT phosphorylase (deod) from *E. coli*"
CDS 59083..59734
FT /tag= v
FT /label= MG050
FT /note= "Previously identified as MORF-20117, the
FT encoded protein shows 83.03 percentage

FT identity to deoxyribose-phosphate aldolase
FT (deoc) from *Mycoplasma pneumoniae*"
CDS complement (64898..65731)
FT /tag= w
FT /label= MG056
FT /note= "Previously identified as MORF-20122, the
FT encoded protein shows 30.25 percent
FT identity to the protein disclosed in
FT GB:D26185_99 from *B. subtilis*"
CDS complement (65713..66249)
FT /tag= x
FT /label= MG057
FT /note= "Previously identified as MORF-20123, the
FT encoded protein shows 38.90 percentage
FT identity to the protein disclosed in
FT GB:D26185_104 from *B. subtilis*"
CDS 81047..82597
FT /tag= y
FT /label= MG067
FT /note= "Previously identified as MORF-19845, the
FT encoded protein shows 28.84 percentage
FT identity to glutamic acid specific protease
FT (spase) from *Staphylococcus aureus*"
CDS 91065..91919
FT /tag= z
FT /label= MG070
FT /note= "Previously identified as MORF-20136, the
FT encoded protein shows 34.8 percentage
FT identity to ribosomal protein S2 (pss2)
FT from *Spirulina plantensis*"
Alignment Scores:
Pred. No.: 2.98e+04 Length: 580073
Score: 42.00 Matches: 9
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.00% Mismatches: 0
Query Match: 85.71% Indels: 0
DB: 18 Gaps: 0
US-09-823-649a-6 (1-11) x AAT58840 (1-580073)
QY 1 LeuSerIysArgIleGlyLeuSerValSer 10
Db 416013 CTGAGTAAACGGTGAAGATTATCAGTTAGT 416042
RESULT 11
ID AAF22306 standard; DNM: 163319 BP.
XX AAF22306/C
AC AAF22306;
XX
DT 20-MAR-2001 (first entry)
XX
DE Arabidopsis thaliana chromosome 4 centromere.
XX
KM Centromere; microsome; vector; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200055325-A2.
XX
PD 21-SEP-2000.
XX
PE 17-MAR-2000; 2000WO-US07392.
XX
PR 18-MAR-1999; 99US-0125219.
XX 01-APR-1999; 99US-0127409.
PR 18-MAR-1999; 99US-0134770.
PR 13-SEP-1999; 99US-0153584.
PR 17-SEP-1999; 99US-0154603.
XX
PA (UYCH-) UNIV CHICAGO.
XX

PI Preuss D, Copenhagen G, Keith K;
 XX WPI: 2000-587529/55.
 DR
 XX
 PT Recombinant DNA construct comprising a plant centromere, useful for
 PT producing stably inherited mitochondria which can serve as vectors for
 PT the construction of transgenic plant and animal cells -
 XX
 PS Claim 79; Page 1389-1451; 1449pp; English.
 CC The present invention relates to a recombinant DNA construct of a plant
 CC (Arabidopsis thaliana) centromere. The constructs are useful for
 CC producing stably inherited mitochondria which can serve as vectors for
 CC the construction of transgenic plant and animal cells expressing
 CC selected proteins such as hormones, enzymes, interleukins, clotting
 CC factors, cytokines, antibodies, and growth factors.
 CC
 SQ Sequence 163319 BP; 53475 A; 30117 C; 28735 G; 50992 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.08e+04 Length: 163319
 Score: 41.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 90.00% Mismatches: 0
 Query Match: 83.67% Indels: 0
 DB: 21 Gaps: 0
 US-09-823-649A-6 (1-11) x AAF22306 (1-163319)
 QY 1 leuserlysarqllleclyleuserlyser 10
 Db 43646 CTTCAACGAAAGGATGCGCTGCGCTGCA 43617
 RESULT 12
 ABA51767/C
 ID ABA51767 standard; DNA; 471 BP.
 XX
 AC ABA51767;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #72.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 XX
 PS Claim 1; SEQ ID NO 72; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at http://wipo.int/pub/published_pat_sequences.
 CC
 SQ Sequence 471 BP; 117 A; 94 C; 104 G; 156 T; 0 other;
 Alignment Scores:
 Pred. No.: 26.9 Length: 471
 Score: 39.00 Matches: 8
 Percent Similarity: 90.91% Conservative: 2
 Best Local Similarity: 72.73% Mismatches: 1
 Query Match: 79.59% Indels: 0
 DB: 22 Gaps: 0
 US-09-823-649A-6 (1-11) x ABA51767 (1-471)
 QY 1 leuserlysarqllleclyleuserlyser 11
 Db 241 TTACMAACGAGGATGCGCTTCAAGTTCAAG 209
 RESULT 13
 ABA21596/C
 ID ABA21596 standard; DNA; 471 BP.
 XX
 AC ABA21596;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Probe #62 for gene expression analysis in human heart cell sample.
 XX
 KW Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 1; SEQ ID NO 62; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosticating diseases of the human heart and vascular system

OY 1 LeuserLysArgIleGlyLeuserValSerGlu 11
| | | : : | | | | | : : | | | | | | | | |
Db 241 TTACAAACGCGCTAGCTCTTCAGTTTCAG 209

Search completed: January 15, 2003, 12:56:51
Job time : 202.143 secs

US-09-823-649A-6 (1-11) x US-09-221-017B-720 (1-541)

QY 1 LeuserlysaArgileGlyLeuserValSer 10
Db 267 CTTCCAAAGCATGGCTCTGATCTTCT 238

RESULT 14

US-09-453-702B-3/C

; Sequence 3, Application US/09453702B

; Patent No. 6365723

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Valerie

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US-09/453,702B

; FILING DATE: 03-Dec-1999

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296,95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 635

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-453-702B-3

Alignment Scores:

Pred. No.: 108

Score: 33.00

Percent Similarity: 88.89%

Best Local Similarity: 88.89%

Query Match: 67.35%

DB: 4

US-09-823-649A-6 (1-11) x US-09-453-702B-3 (1-635)

QY 2 SerlysaArgileGlyLeuserValSer 10

Db 485 AGTAAAGATGGGTTAAGCCGTTCT 459

RESULT 15

US-09-199-637A-26/C

; Sequence 26, Application US/09199637A

; Patent No. 6355411

; GENERAL INFORMATION:

APPLICANT: Ausubel, Frederick

APPLICANT: Goodman, Howard M.

APPLICANT: Rahme, Laurence G.

APPLICANT: Mahajan-Miklos, Shalina

APPLICANT: Tan, Man-Wah

APPLICANT: Cao, Hui

APPLICANT: Drenkard, Eliana

APPLICANT: Tsongalis, John

TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID

FILE REFERENCE: 00786/361002

CURRENT APPLICATION NUMBER: US/09/199,637A

PRIOR FILING DATE: 1998-11-25

PRIOR APPLICATION NUMBER: 60/066,517

PRIOR FILING DATE: 1997-11-25

NUMBER OF SEQ ID NOS: 437

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 26

LENGTH: 642

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-199-637A-26

Alignment Scores:

Pred. No.: 110

Score: 33.00

Percent Similarity: 88.89%

Best Local Similarity: 77.78%

Query Match: 67.35%

DB: 4

US-09-823-649A-6 (1-11) x US-09-199-637A-26 (1-642)

QY 1 LeuserlysaArgileGlyLeuserVal 9

Db 371 TTAGCAAGCGCTTGGCTCTCGGTC 345

Search completed: January 15, 2003, 12:58:54

Job time : 31 secs

APPLICANT: Crossland, Lyle D
APPLICANT: Harper, Stacy M
TITLE OF INVENTION: Method of Hybrid Seed Production Using
TITLE OF INVENTION: Conditional Female Sterility
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS: 14
ADDRESSEE: No. 639213artis Corporation - Patent & Trademark
ADDRESSEE: Dept.
STREET: P.O. Box 12257
CITY: Research Triangle Park
STATE: NCNY
COUNTRY: USA
ZIP: 22057
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/575,602
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/032,086
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1915/Reg
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 6596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Promoter
LOCATION: 1..3790
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "5' Regulatory Region of
OTHER INFORMATION: B20014-2"
OTHER INFORMATION: /evidence= EXPERIMENTAL
FEATURE:
NAME/KEY: misc_signal
LOCATION: 4427..6397
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "3' Regulatory Region
OTHER INFORMATION: for B20014-2"
OTHER INFORMATION: /evidence= EXPERIMENTAL
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3789..3791
OTHER INFORMATION: /function= "ATG translation start"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4402..4404
OTHER INFORMATION: /function= "translation stop"
US-09-575-602-11
Alignment Scores:
Pred. No.: 1.17e+03 Length: 6596
Score: 34.00 Matches: 8
Percent Similarity: 81.82% Conservative: 1
Best Local Similarity: 72.73% Mismatches: 2
Query Match: 69.39% Indels: 0
DB: 4 Gaps: 0
US-09-823-649a-6 (1-11) x US-09-575-602-11 (1-6596)

Qy 1 LeuserLysArg11eGlyLeuserValSerGlu 11
|||||
Db 1175 CTTTCGAAGTAAGGATTCATTTCTCGAG 1207
RESULT 13
US-09-221-017B-720/c
Sequence 720, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Montoy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 720:
SEQUENCE CHARACTERISTICS:
LENGTH: 541 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...541
US-09-221-017B-720
Alignment Scores:
Pred. No.: 88.7 Length: 541
Score: 33.00 Matches: 7
Percent Similarity: 90.00% Conservative: 2
Best Local Similarity: 70.00% Mismatches: 1
Query Match: 67.35% Indels: 0
DB: 4 Gaps: 0

LENGTH: 5406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-166

Alignment Scores:

Pred. No.:	321	Length:	5406
Score:	36.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	4
Best Local Similarity:	60.00%	Mismatches:	0
Query Match:	73.47%	Indels:	0
DB:	4	Gaps:	0

US-09-823-649A-6 (1-11) x US-08-961-527-166 (1-5406)

QY 2 SerlysergilegylLeuserValserglu 11

DB 2152 TTCAGAGACCTGGCTATCTATTAAAGCA 2123

RESULT 10

US-08-961-527-181
; Sequence 181, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 181:

SEQUENCE CHARACTERISTICS:

LENGTH: 8651 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

Alignment Scores:

Pred. No.:	969	Length:	8651
Score:	35.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	71.43%	Indels:	0
DB:	4	Gaps:	0

US-09-823-649A-6 (1-11) x US-08-961-527-181 (1-8651)

QY 1 LeuserlysergilegylLeuserValserglu 11

DB 2129 TTAGCAAAAATFTGGTATTCTGTACAGAA 2161
|||||

RESULT 11

US-08-905-223-39
; Sequence 39, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Duclercq, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Win95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/905,223

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 427 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: DOUBLE

TOPOLOGY: LINEAR

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

ORGANISM: Homo Sapiens

TISSUE TYPE: Brain

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 191..288

IDENTIFICATION METHOD: Von Heijne matrix

OTHER INFORMATION: score 10.8

US-08-905-223-39

Alignment Scores:

Pred. No.:	39.2	Length:	427
Score:	34.00	Matches:	7
Percent Similarity:	81.82%	Conservative:	2
Best Local Similarity:	63.64%	Mismatches:	2
Query Match:	69.39%	Indels:	0
DB:	4	Gaps:	0

US-09-823-649A-6 (1-11) x US-08-905-223-39 (1-427)

QY 1 LeuserlysergilegylLeuserValserglu 11

DB 332 TTCAGAGACCTGGCTATCTATTAAAGCA 364

RESULT 12

US-09-575-602-11
; Sequence 11, Application US/09575602
; Patent No. 6392123
; GENERAL INFORMATION:

ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/987,151
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11748
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1785 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-987-151-3

Alignment Scores:
Pred. No.: 81.2 Length: 1785
Score: 36.00 Matches: 6
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 60.008 Mismatches: 0
Query Match: 73.478 Indels: 0
Gaps: 0

US-09-823-649a-6 (1-11) x US-08-987-151-3 (1-1785)

QY 2 SerLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||

Db 1065 DCCAGAGACUUGGCCUADUUAACGAA 1094

RESULT 8
US-08-687-590-61
Sequence 61, Application US/08687590
Patent No. 6255070
GENERAL INFORMATION:
APPLICANT: Willison, Keith Robert
APPLICANT: Kubota, Hiroshi
APPLICANT: Ashworth, Alan
TITLE OF INVENTION: Folding Proteins
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,590
FILING DATE: 31-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/00192

FILING DATE: 31-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9401791.0
FILING DATE: 31-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9418234.2
FILING DATE: 09-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 084619-000000US
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 1853 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 86..1732
US-08-687-590-61

Alignment Scores:
Pred. No.: 85 Length: 1853
Score: 36.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.47% Indels: 0
Gaps: 0

US-09-823-649a-6 (1-11) x US-08-687-590-61 (1-1853)

QY 4 ArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||

Db 443 AGGATTGGCTGTCAATATCAGAG 466

RESULT 9
US-08-961-527-166/c
Sequence 166, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: PB340P1
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 166:
SEQUENCE CHARACTERISTICS:

Query Match: 77.55% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-6 (1-11) x US-09-134-001C-2379 (1-495)

QY 1 LeuSerLysArgIleGlyLeuSerValSer 10
|||||:|||||:|||||:|||||:|||||
DB 398 C|ATCCAGAGCTGATGGCTATCTATTAAACGAA 369

RESULT 5
US-08-987-151-4
; Sequence 4, Application US/08987151
; Patent No. 6162617
; GENERAL INFORMATION:
; APPLICANT: Jaskunas Jr., Stanley R.
; APPLICANT: Zhao, Genshi
; APPLICANT: Peery, Robert B.
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Rosteck Jr., Paul R.
; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence
; TITLE OF INVENTION: diag
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,151
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-987-151-4

Alignment Scores:
Pred. No.: 71 Length: 1602
Score: 36.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 60.00% Mismatches: 0
Query Match: 73.47% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-6 (1-11) x US-08-987-151-4 (1-1602)

QY 2 SerLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||
DB 1158 TCCAGAGAGCTGGCTATCTATTAAACGAA 1187

RESULT 6
US-08-987-151-1
; Sequence 1, Application US/08987151
; Patent No. 6162617

GENERAL INFORMATION:
; APPLICANT: Jaskunas Jr., Stanley R.
; APPLICANT: Zhao, Genshi
; APPLICANT: Peery, Robert B.
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Rosteck Jr., Paul R.
; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence
; TITLE OF INVENTION: diag
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,151
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1785 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 1..1782
US-08-987-151-1

Alignment Scores:
Pred. No.: 81.2 Length: 1785
Score: 36.00 Matches: 6
Percent Similarity: 100.00% Conservative: 4
Best Local Similarity: 60.00% Mismatches: 0
Query Match: 73.47% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-6 (1-11) x US-08-987-151-1 (1-1785)

QY 2 SerLysArgIleGlyLeuSerValSerGlu 11
|||||:|||||:|||||:|||||:|||||
DB 1065 TCCAGAGAGCTGGCTATCTATTAAACGAA 1094

RESULT 7
US-08-987-151-3
; Sequence 3, Application US/08987151
; Patent No. 6162617
; GENERAL INFORMATION:
; APPLICANT: Jaskunas Jr., Stanley R.
; APPLICANT: Zhao, Genshi
; APPLICANT: Peery, Robert B.
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Rosteck Jr., Paul R.
; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence
; TITLE OF INVENTION: diag
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

ORGANISM: Thermophilus africanus
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2676
US-08-458-819-11
Alignment Scores:
Pred. No.: 0.15 Length: 2679
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-823-649a-6 (1-11) x US-08-458-819-11 (1-2679)
OY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db 2218 CTTTCAAGAGAAATGGCTTACTGTTTCAGAG 2250
RESULT 3
PCT-US91-07035-11
Sequence 11, Application PC/TUS9107035
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cetus Corporation
STREET: 1400 Fifty-third Street
CITY: Emeryville
STATE: California
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/07035
FILING DATE: 19910930
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471

FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stas Ph.D, Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: Case No. 2580
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-420-3300
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2679 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Thermophilus africanus
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2676
PCT-US91-07035-11
Alignment Scores:
Pred. No.: 0.15 Length: 2679
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0
US-09-823-649a-6 (1-11) x PCT-US91-07035-11 (1-2679)
OY 1 LeuSerLysArgIleGlyLeuSerValSerGlu 11
Db 2218 CTTTCAAGAGAAATGGCTTACTGTTTCAGAG 2250
RESULT 4
US-09-134-001C-2379/c
Sequence 2379, Application US/09134001C
Patent No. 6380470
GENERAL INFORMATION:
APPLICANT: Lynn Doucelle-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPY
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2379
LENGTH: 495
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2379
Alignment Scores:
Pred. No.: 5.82 Length: 495
Score: 38.00 Matches: 8
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 80.00% Mismatches: 0

;; FILING DATE: 15-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 143,441
;; FILING DATE: 12-JAN-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 063,509
;; FILING DATE: 17-JUN-1987
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 899,241
;; FILING DATE: 22-AUG-1986
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 746,121
;; FILING DATE: 15-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US90/07641
;; FILING DATE: 21-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 585,471
;; FILING DATE: 20-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 455,611
;; FILING DATE: 22-DEC-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 609,157
;; FILING DATE: 02-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 557,517
;; FILING DATE: 24-JUL-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Luann Cseert
;; REGISTRATION NUMBER: 31,822
;; REFERENCE/DOCKET NUMBER: Case No. 546591 8753
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 814-2972
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2679 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Thermophilus africanus
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..2676
;; US-07-977-434-11

Alignment Scores:
Pred. No.: 0.15 Length: 2679
Score: 49.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 1 Gaps: 0

US-09-823-649a-6 (1-11) x US-07-977-434-11 (1-2679)
Qy 1 leuserlysaqllleglyleuserlysaqll 11
Db 2218 CTTCAAGAGAAATGTGTTGTTCAAG 2250

RESULT 2
US-08-458-819-11
; Sequence 11, Application US/08458819
; Patent No. 5795762
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David H.
; APPLICANT: Abramson, Richard D.
; TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES

;; NUMBER OF SEQUENCES: 38
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hoffmann-La Roche Inc.
;; STREET: 340 Kingsland Street
;; CITY: Nutley
;; STATE: New Jersey
;; ZIP: 07110-1199
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Macintosh
;; OPERATING SYSTEM: 7
;; SOFTWARE: wordperfect 2.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/458,819
;; FILING DATE: 02-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/977,434
;; FILING DATE: 23-FEB-1993
;; APPLICATION NUMBER: US 590,490
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 590,466
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 590,213
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 523,394
;; FILING DATE: 15-MAY-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 143,441
;; FILING DATE: 12-JAN-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 063,509
;; FILING DATE: 17-JUN-1987
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 899,241
;; FILING DATE: 22-AUG-1986
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 746,121
;; FILING DATE: 15-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US90/07641
;; FILING DATE: 21-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 585,471
;; FILING DATE: 20-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 455,611
;; FILING DATE: 22-DEC-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 609,157
;; FILING DATE: 02-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 557,517
;; FILING DATE: 24-JUL-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Luann Cseert
;; REGISTRATION NUMBER: 31,822
;; REFERENCE/DOCKET NUMBER: Case No. 5795762 8753
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 814-2972
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2679 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 12:36:21 ; Search time 26 Seconds
(without alignments)
129.748 Million cell updates/sec

Title: US-09-823-649a-6
Perfect score: 49
Sequence: 1 LSKRIGLSYSE 11

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -f -x
-MODE=frame_plus_p2n_model -DEV=slh
-O=/cgn2_1/USPTO_spool/US09823649.runat.14012003.151002.29127/app_query_fastea_1.1393
-DB=Issued_Patents_NA -QPM=fastap -SUFFIX=rni -MIMATCH=0.1 -DOPC=0
-LOOEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=bits -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR=SCORE=100 -THR=MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.rni.1.35 &runat.14012003.151002.29127 -MCP=6 -ICPU=3
-NO_XLIFY -NO_MMAP -LARGEROBJ -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MARK_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA:
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCUTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	49	100.0	2679	1	US-07-977-434-11	Sequence 11, Appl
2	49	100.0	2679	1	US-08-458-819-11	Sequence 11, Appl
3	49	100.0	2679	5	PCR-US91-07035-11	Sequence 11, Appl
4	38	77.6	495	4	US-09-134-001C-2379	Sequence 2379, Ap
5	36	73.5	1602	4	US-08-987-151-4	Sequence 4, Appl
6	36	73.5	1785	4	US-08-987-151-1	Sequence 1, Appl
7	36	73.5	1785	4	US-08-987-151-3	Sequence 3, Appl
8	36	73.5	1853	4	US-08-687-590-61	Sequence 61, Appl
9	36	73.5	5406	4	US-08-961-527-166	Sequence 166, App
10	35	71.4	8651	4	US-08-961-527-181	Sequence 181, App
11	34	69.4	427	4	US-08-905-223-39	Sequence 39, Appl
12	34	69.4	6596	4	US-09-575-602-11	Sequence 11, Appl

C 13	33	67.3	541	4	US-09-221-017B-720	Sequence 720, App
C 14	33	67.3	635	4	US-09-453-702B-3	Sequence 3, Appl
C 15	33	67.3	642	4	US-09-199-637A-26	Sequence 26, Appl
C 16	33	67.3	685	4	US-09-221-017B-861	Sequence 861, App
C 17	33	67.3	957	4	US-09-199-637A-22	Sequence 22, Appl
C 18	33	67.3	1050	3	US-08-755-587-24	Sequence 24, Appl
C 19	33	67.3	1346	2	US-08-557-128-2	Sequence 2, Appl
C 20	33	67.3	1407	4	US-09-199-637A-28	Sequence 28, Appl
C 21	33	67.3	2769	4	US-09-118-408-1	Sequence 1, Appl
C 22	33	67.3	2769	4	US-09-506-855-1	Sequence 1, Appl
C 23	33	67.3	3141	4	US-09-199-637A-12	Sequence 12, Appl
C 24	33	67.3	8533	1	US-07-846-181-6	Sequence 6, Appl
C 25	33	67.3	8533	1	US-07-845-989-6	Sequence 6, Appl
C 26	33	67.3	9244	4	US-08-961-527-68	Sequence 68, Appl
C 27	33	67.3	10223	4	US-08-961-527-73	Sequence 73, Appl
C 28	33	67.3	42235	4	US-09-199-637A-1	Sequence 1, Appl
C 29	32	65.3	87	4	US-09-364-543-98	Sequence 98, Appl
C 30	32	65.3	400	4	US-09-020-956-76	Sequence 76, Appl
C 31	32	65.3	400	4	US-09-030-607-76	Sequence 76, Appl
C 32	32	65.3	400	4	US-09-605-785-76	Sequence 76, Appl
C 33	32	65.3	400	4	US-09-439-313-76	Sequence 76, Appl
C 34	32	65.3	400	4	US-09-352-616A-76	Sequence 76, Appl
C 35	32	65.3	400	4	US-09-232-149A-76	Sequence 76, Appl
C 36	32	65.3	775	4	US-09-020-956-25	Sequence 25, Appl
C 37	32	65.3	775	4	US-09-030-607-25	Sequence 25, Appl
C 38	32	65.3	775	4	US-09-605-785-25	Sequence 25, Appl
C 39	32	65.3	775	4	US-09-439-313-25	Sequence 25, Appl
C 40	32	65.3	775	4	US-09-352-616A-25	Sequence 25, Appl
C 41	32	65.3	775	4	US-09-232-149A-25	Sequence 25, Appl
C 42	32	65.3	999	4	US-08-870-511-7	Sequence 7, Appl
C 43	32	65.3	999	4	US-08-870-511-9	Sequence 9, Appl
C 44	32	65.3	999	4	US-08-870-511-11	Sequence 11, Appl
C 45	32	65.3	1160	1	US-08-290-448A-77	Sequence 77, Appl

ALIGNMENTS

RESULT 1
US-07-977-434-11
Sequence 11, Application US/07977434
Patent No. 546591
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSER: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: WordPerfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 16, 2003, 08:03:43 ; Search time 1525.14 Seconds
(Without alignments)

209,902 Million cell updates/sec

Title: US-09-823-649a-6
Perfect score: 49
Sequence: 1 LSKRIGLSYSE 11

Scoring table: BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -
-MODE=frame+p2n model -DEV=tlh
-Q/cgn2_1/USPO.spool/US09823649/runat_14012003.151002.29139/app.query.fasta_1.1393
-DB=GenEmbl -QMT=Fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -V=rm-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09823649.ecgn.1.1.3378 &runat_14012003.151002.29139 -NCPU=6 -ICPU=3
-NO.XIPXY -NO.MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.cm:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pin:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htg.hum:*
40: em.htg.mus:*
41: em.htg.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	2679	6 AR023939	AR023939 Sequence
2	49	100.0	2679	6 I15440	I15440 Sequence 11
3	42	85.7	12486	1 U39714	U39714 Mycoplasma
4	42	85.7	147168	2 AC124922	AC124922 Rattus no
5	41	83.7	1245	3 HE088626	HE088626 Helicobacter
6	41	83.7	10673	1 U67546	U67546 Methanococcus
7	41	83.7	47310	6 AX059549	AX059549 Sequence
8	41	83.7	63292	6 F14616	F14616 Arabidopsis
9	41	83.7	120885	2 AC113735	AC113735 Rattus no
10	41	83.7	185133	2 AC095167	AC095167 Rattus no
11	41	83.7	189300	2 AC099720	AC099720 Mus muscu
12	41	83.7	198176	8 ANCHRI17	ANCHRI17 Mus muscu
13	40	81.6	150018	2 AC114382	AC114382 Rattus no
14	40	81.6	170063	2 AC107106	AC107106 Rattus no
15	40	81.6	180674	2 AC094431	AC094431 Rattus no
16	40	81.6	189982	2 AC095409	AC095409 Rattus no
17	40	81.6	200475	2 AC095431	AC095431 Rattus no
18	39	79.6	1609	9 AK098608	AK098608 Homo sapi
19	39	79.6	2267	9 AB047956	AB047956 Homo sapi
20	39	79.6	6394	1 AF311944	AF311944 Methanopy
21	39	79.6	10527	1 AE010436	AE010436 Methanopy
22	39	79.6	76164	2 AC130382	AC130382 Homo sapi
23	39	79.6	113688	9 AC005301	AC005301 Homo sapi
24	39	79.6	124823	9 AC007064	AC007064 Homo sapi
25	39	79.6	226650	2 AL772343	AL772343 Mus muscu
26	38	77.6	248361	2 AC124579	AC124579 Mus muscu
27	38	77.6	468	6 AX141621	AX141621 Sequence
28	38	77.6	1380	1 RACELB	RACELB Sequence
29	38	77.6	3472	1 AF269429	AF269429 Rattus no
30	38	77.6	3472	6 AX144749	AX144749 Sequence
31	38	77.6	4020	3 AY063158	AY063158 Rattus no
32	38	77.6	4105	1 AF269437	AF269437 Rattus no
33	38	77.6	4105	6 AX144757	AX144757 Sequence
34	38	77.6	24171	3 CE031A11	CE031A11 Caenorhabdi
35	38	77.6	71198	2 AC097762	AC097762 Rattus no
36	38	77.6	88647	2 AC114476	AC114476 Magnapor
37	38	77.6	142321	2 AC120090	AC120090 Rattus no
38	38	77.6	144255	2 AC107109	AC107109 Rattus no
39	38	77.6	150576	2 AC119098	AC119098 Rattus no
40	38	77.6	152763	2 AC119715	AC119715 Rattus no
41	38	77.6	158346	2 AC098756	AC098756 Rattus no
42	38	77.6	160222	2 AC098204	AC098204 Rattus no
43	38	77.6	160580	2 AC120975	AC120975 Rattus no
44	38	77.6	165032	2 AC097597	AC097597 Rattus no
45	38	77.6	185131	2 AC127739	AC127739 Rattus no

RESULT 1

ALIGNMENTS

KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
		Mycoplasma genitalium.						
		Mycoplasma genitalium						
		Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.						
		1 (bases 1 to 12486)						
		Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A.,						
		Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.,						
		Kelly,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M.,						
		Fuhrmann,J.L., Nguyen,D.T., Uterback,T., Saudek,D.M., Phillips,C.A.,						
		Merrick,J.M., Tomb,J., Dougherty,B.A., Bott,K.F., Hu,P.C.,						
		Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.						
		The minimal gene complement of Mycoplasma genitalium						
		Science 270 (5235), 397-403 (1995)						
		96026346						
		7569993						
		2 (bases 1 to 12486)						
		Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A.,						
		Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.,						
		Kelly,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M.,						
		Fuhrmann,J.L., Nguyen,D.T., Uterback,T., Saudek,D.M., Phillips,C.A.,						
		Merrick,J.M., Tomb,J., Dougherty,B.A., Bott,K.F., Hu,P.C.,						
		Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.						
		Direct Submission						
		Submitted (29-OCT-1995) The Institute for Genomic Research, 9712						
		Medical Center Drive, Rockville, MD 20850, USA						
		3 (bases 1 to 12486)						
		Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A.,						
		Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.,						
		Kelly,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M.,						
		Fuhrmann,J.L., Nguyen,D.T., Uterback,T., Saudek,D.M., Phillips,C.A.,						
		Merrick,J.M., Tomb,J., Dougherty,B.A., Bott,K.F., Hu,P.C.,						
		Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.						
		Direct Submission						
		Submitted (19-OCT-1998) The Institute for Genomic Research, 9712						
		Medical Center Dr. Rockville, MD 20850, USA						
		On Nov 5, 1998 this sequence version replaced gi:1046022.						
		Location/Qualifiers						
		1..12486						
		/organism="Mycoplasma genitalium"						
		/isolate="G37"						
		/db_xref="taxon:2097"						
		complement(14..901)						
		/gene="MG326"						
		complement(14..901)						
		/gene="MG327"						
		/note="similar to GB:U00089 SP:P75312 PID:1674048 percent						
		identity: 68.94; identified by sequence similarity;						
		putative"						
		/codon_start=1						
		/transl_table=4						
		/product="conserved hypothetical protein"						
		/protein_id="AAC71550.1"						
		/db_xref="gi:3844906"						
		/translation="MKKTAITTPSPASIRGKINGVYILPQIVYDGKSPDGIETIDH						
		YDVHFLKLNKHGHNSTSLSPROSLKIFPEKTRVDFPILPLSKSLSTDMY						
		QIAKLSBNKDKKEIVPETSDAISLKLVDIALVDKGCNQTAKKVESKQNI						
		LSAVTKRNVQRRKGRISGLKFTITLLRVAPILDPKCVNIGAKVSPQAVETKI						
		EGVKKRFGDNKIKRIGFCYSCKNVANEIKKTIITPEKNIINQNEIEMAFSTSI						
		IIVHTGIDAFSISILINK"						
		com						

```

gene
CDS
    /transl_table=4
    /product="conserved hypothetical protein"
    /protein_id="AAC71552.1"
    /db_xref="GI:3844908"
    /translation="MAYDKBLEISDFENDELDEKTLKELVQRTNNILSPSKITLIP
    ERNLTEFFETVDIAEKESIVSFNMWIDLVKQDKWPNVNLNVAQQLTREFE
    QVQDTMAFQEDVDQSVLKEIKTGQELKPSVITAEEDDDEIKPATQVSEFLN
    OPSEINETKKPEVQIESTDKVKEPEQPDQVSIENLTKAINPHKTIQYDQNDQPF
    VYKRIKEQHPETKVELDQYNNKELLENDLKQIDDKENNDQIDLEOEDL
    KRIISEKSKTHKRIKODDLOENNDLEOLOKPAVNPISDEVNELLELKOGEA
    LLSDDDLAKKSSNVQOQALIPVNNQINELQNLITREANQDLDBENDELK
    NELKLDHNTSDNEKEYDQDLNQYELPDENETKFKLDYKQOALNDQYRTISALK
    HENDVLDELIMTFRSKDNEFNNTKNSFEQKALDEKLNGLTQONQLODKLAELEW
    NEKSNLNTQNLVNLQOLKDSQNLSEVNAQDKLALLESVNLANKINDLEDELSGE
    NSNNLTLAQDEHETLOESTYKLTDEFEKKNKINDANEDQYDQLSAEPEFNSLEK
    AKOSLSASDSENNOLKQOINSELENAKLEOFPTVSDHEIDLETIKELFENQ
    ALQNOLOTFNDISANGTEETKEKNSDEKVEIKKPRIKKDPYIQKKDKLAKSKRE
    RIQAYAEKRLKINANE"
    complement(4016..5362)
    /gene="MG329"
    /complement(4016..5362)
    /note="similar to GB:U00021 PID:467147 percent identity:
    32.18; identified by sequence similarity; putative"
    /codon_start=1
    /transl_table=4
    /product="conserved hypothetical GTP-binding protein"
    /protein_id="AAC71553.1"
    /db_xref="GI:3844909"
    /translation="WFTVALIGRTNVGKSTLFLNLQKPMALVSDIPNTTRIRIGIC
    EMIKKRIAFIDTGLIAKQPIQLALQVQAALQSAKALIFLVLSQEDLNDDEYVA
    KYKRNKDKPVLIVVKNKNE-NKTAEBTLKDYSLGGRVYISAHIGIGDMLD
    LKYLQKQDLPNENNDLAKIRFCVIGKPNVSKSLNOLYKONRVLVSNSSGKTRPID
    VPLATNGKRPILLIDAGIKRKGKINGGIEASIKTKLALARSVILLAWDSKRISE
    QBEVIGIQAQALIPVITIIYKMDLVTKNNNTNAYKKKLIKHPKHLDPAPLFTSVL
    KNORLNTTFEQKTILOSQLETKVATPLNDVYQQAQLYNQPPIPKKGRQITTYAQTK
    SQIPPEVLEFNDPRLHESYARFLNKIKRNGFNSVPISLFKSKNARIKRPV"
    complement(5366..6019)
    /gene="MG330"
    /complement(5366..6019)
    /note="similar to GB:U00089 SP:P75308 PID:1674043 percent
    identity: 65.42; identified by sequence similarity;
    putative"
    /codon_start=1
    /transl_table=4
    /product="cytidylate kinase (cmk)"
    /protein_id="AAC71554.1"
    /db_xref="GI:3844910"
    /translation="MNNQIALIDGSSSSGSSVAKKIAHELDYFSSGKWRAPAYVM
    QVNRLLNDLFLKINQINMRFERDAYVNNADITVITQSVANAKSLAVPNNRKI
    AVYKQKLAENKINVDGRDIGVYVKNQKLEFEDAVLEKQRIQRLQMDGSLSNK
    KIKELIOELKORDIDSSRTADPLKKAQAIYLDTSLSFDVAVYKQLKKAQVYKL"
    complement(6028..6666)
    /gene="MG331"
    /complement(6028..6666)
    /note="similar to GB:U00089 SP:P75307 PID:1674042 percent
    identity: 39.34; identified by sequence similarity;
    putative"
    /codon_start=1

gene
CDS
    /transl_table=4
    /product="conserved hypothetical protein"
    /protein_id="AAC71555.1"
    /db_xref="GI:3844911"
    /translation="MGVEKEFRFROSDENKIKKALINAQNTESMKQKJLNK
    ILLNHPSEPNKVPVHHTPEPKLPTLQELVDLNDTPKILKEKVRMMNAATN
    QRSSGYESWISDDKGLIKIINLSKPEANQKQKNSNLSQDKYNNLIYVSTES
    FAPKSIDNIEPNLFNANKALFKNVILVILISPMLSILFPLILVSPVSV"
    complement(6654..7373)
    /gene="MG332"
    /complement(6654..7373)
    /note="similar to GB:U00089 SP:P75306 PID:1674041 percent
    identity: 75.64; identified by sequence similarity;
    putative"
    /codon_start=1
    /transl_table=4
    /product="conserved hypothetical protein"
    /protein_id="AAC71556.1"
    /db_xref="GI:3844912"
    /translation="MPKRLIANTQPKQTSKQLOKLAKRIASAVKGGTINQSNP
    HLKAVDALAKGLSMDIKRNTGSEKQETISRECYEIPGPGVGIIVGLINPN
    KLLSLNGYIAKLGKGLAKGAKNSYKIKK-3GGITPNKNNYKFDLIELLIDNINLID
    VYDECEPISHSNSTFAKELKLNKNSIYDSIKIYPLIYVDLQKQOTLISRL
    NACEEDDIDQVVAHNPWEE"
    complement(7380..7974)
    /gene="MG333"
    /note="this region contains an authentic frame shift and
    is not the result of a sequencing artifact; similar to
    GB:U00089 SP:P75305 PID:1674040 percent identity: 74.75;
    identified by sequence similarity; putative"
    /complement(8011..10524)
    /gene="MG334"
    /complement(8011..10524)
    /note="similar to GB:U00089 SP:P75304 PID:1674039 percent
    identity: 63.96; identified by sequence similarity;
    putative"
    /codon_start=1
    /transl_table=4
    /product="valyl-tRNA synthetase (vals)"
    /protein_id="AAC71557.1"
    /db_xref="GI:3844913"
    /translation="MKDKFSQKNYDNLVSDGLYEIVNNAGFKPKDKKNSVTAIIP
    PPLTLTLHIGAEVSGITQDIRFKMGQFQINIPGFDIAGATQTKETKIAKEN
    QKLYEDDKSEMIIMNNALNOSEIITKQKSLGVCILNMSKFTLSQANKIVNQC
    KNLYEGEYGAQYTLVMDRKLNTATISNIEVKNVGNHLYVYKLIANSKQELIVA
    TTRPEIIPADQCLANPKDKRTYNKKNLYVNPILQKQIPIYVTSYVDIKKGIILKC
    TPADHNDYRIINTYKRFQPLSCIDSGTLNQAASKFQGLSVIQRANKIVAKLKNKL
    VKSIPLTNSVCHSRSCIVVEPMLSKQWVVDLPKLDHLYLKKYDPPIPKRPNKQVSN

```

Alignment Scores:

Pred. No.:	113	Length:	12486
Score:	42.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	90.00%	Mismatches:	0
Query Match:	85.71%	Indels:	0
DB:	1	Gaps:	0

US-09-823-649a-6 (1-11) x U39714 (1-12486)

QY 1 LeuSerIysArgIleGlyLeuSerValSer 10

|||||

Db 7057 CTGAGTAACGGTAGATATACAGTTAGT 7086

|||||

RESULT 4 AC124922/c 147168 bp DNA linear HTG 24-JUL-2002

LOCUS AC124922 Rattus norvegicus clone CH230-397L24, *** SEQUENCING IN PROGRESS

DEFINITION AC124922 *** 28 unordered pieces.

ACCESSION AC124922

VERSION AC124922.2 GI:21671429

KEYWORDS HTG; HTGS_PHASE1.

gene
CDS
/gene="MJ1033"
1191..2144
/note="similar to GB:142023 SP:P4498 P1D:1006270
P1D:1205285 P1D:1221151 percent identity: 25.33;
identified by sequence similarity; putative"
/transl_table=11
/product="magnesium and cobalt transport protein (cora)"
/protein_id="AAB99037.1"
/db_xref="GI:1499876"
/translation="MTVIAIADGSIIVEPKIDELISFEDRYLWIDCPKDELEKTL
SKRIGISVSDLOIGDEDEIPREDEDEYLIYKAPLEEDITTSIGIYKNNLL
TTHSKIKAGIKGILKISTKPRIVPERGIGLILHILNETRSYILNLEDELE
LEKILAGIDRWKEXIKIGIKRTIVFHSILANDVILKIKVILPITTKRENEE
DITDTIDIDMSATIREVLSMDITLSLEIKNOKIMKILIVVITIRAPWMLTIGI
YGNFESYLPANNPQGFVLMALMVIIIMFYIFRRSGWI"
2154..2417
/gene="MJ1034"
2154..2417
/note="similar to GB:X12791 SP:P09132 P1D:36113 percent
identity: 32.89; identified by sequence similarity;
putative"
/transl_table=11
/product="signal recognition particle, subunit SRP19"
/protein_id="AAB99038.1"
/db_xref="GI:1499877"
/translation="MIMPSTIDKRSRREGKRVPEELIKPSLKDIEKALKIGLE
PKIYRDRPRQHEICGCVDEYKGNKQLLEIKIKGN"
/complement(2440..3273)
/gene="MJ1035"
/complement(2440..3273)
/note="similar to P1D:685181 P1D:726050 SP:P55300
GB:AEO00666 percent identity: 66.43; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="N5, N10-methylene-tetrahydromethanopterin
dehydrogenase (mtd)"
/protein_id="AAB99039.1"
/db_xref="GI:1591688"
/translation="MWKIGIKGNGISPVVDALDERARDKIDAVRIGSGAKD
PESVPEVTKWKEEYKDPFIYIGPNAPGKRRKEIISOGIPIVITDAGCIPVK
DEWQGGIYIILKCDPMIGARRERLDPYKMAFNADVIRIYLAQTGRLTQVQALDKM
IDAKKGEKIELEPKIVITEQNAVEAFETPPAKAKAAFTIAEKVGDVAVGCFMT
KEAEKIPIVASAHEMIYAAKIVDEARELEKAMDAYSRRKHHEGRKLSKALMEKP
E"
3682..4383
/gene="MJ1036"
3682..4383
/note="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=11
/product="M. jannaschii predicted coding region MJ1036"
/protein_id="AAB99045.1"
/db_xref="GI:1591689"
/translation="MSNNEVPFIFIALISFVWCIGELIAYSVLSLKYKEPAISFG
FIFGVALIILPKSYSNMFVLYLGMIVYILEKVLACPLSKKYCEVDNLEENRI
KETIPISEPIHPIIDGLIIVAYSEISGLPIALIMKHLPGAGVILSPKGYKQML
YKGFVSECVIGTGVLTAKDVSITKLILASGCVFLGAFELMLAPHIYHKEKEKFL
YILGILVIGITALH"
4446..5660
/gene="MJ1037"
4446..5660
/note="similar to P1D:1054861 SP:P50909 percent identity:
64.12; identified by sequence similarity; putative"
/codon_start=1

gene
CDS
/transl_table=11
/product="tryptophan synthase beta subunit (trpB)"
/protein_id="AAB99040.1"
/db_xref="GI:1591690"
/translation="WSILIKYKDMYPDENKRGTYIGKAVPEPIIMPAIAHLHPKRRP
MINNEGNFEESEYALRDYVGRPTPIYAEARESEELGCKYKAKEDIAHGHKINNA
LGQALLAKMKGRKRYIAETGAGOHVATAACAKIGECITIMAKADVEKOKLNVFRA
ELMGAKVPIVPEGSSQTLKDAVENALDMNTYRTYITVYLSGASLPHYPMVREORY
ICKELKEQILEKESGLPIVYACVCGGSNAITAFEPFLDDVRLYANFAGGCGITGCM
HNASLDCGEVGVLMGAKIYNKEDFQIEESTISAGLDYGVGPBISPLKIDCRKA
VCVYDEALEAPOLLCRLEGILPALSSHALVAVALADKLDKDDWVIMISGDDKD
WQVAVAKALREI"
5837..6691
/gene="MJ1038"
5837..6691
/note="similar to GB:M65060 SP:P26920 P1D:149750
GB:AEO00666 percent identity: 47.77; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="tryptophan synthase alpha subunit (trpA)"
/protein_id="AAB99041.1"
/db_xref="GI:1591691"
/translation="MEMMINKIAKEFEELKNGKAKAFYAVYGGDPNLEISERALEVI
CKHADIVEIGPESDPVADGITIOKADVRLNLSGNMPLKAKELAKINEKAPVAVF
LHYVNIIFKMGEEFVKKACAGSGIIVPDLPIEADSLYVCKGQNDLPIVAPV
TPHRIKILEKSGSFYVYVSGITGAREKAEETKELIRVYKPKRIACYGRGTS
KRHVEITEITADGAIVSALVIXERHLDENGQIKDEKFLKELEFVKNLKEGYSK
KAVYALKN"
6866..7102
/gene="MJ1039"
6866..7102
/note="similar to GB:X59151 SP:P41559 P1D:561387
P1D:809732 percent identity: 49.33; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, subunit H (trpH)"
/protein_id="AAB99042.1"
/db_xref="GI:1591692"
/translation="MKVTDHILVKEHIVKEKEVEHILKRYNIRIKQOLPPIYEDDEVI
QEIGAKEDGVVRVIRKSPYAGVSAYRLVIRKII"
7110..8606
/gene="MJ1040"
7110..8606
/note="similar to GB:X73293 P1D:505290 SP:P41558 percent
identity: 70.99; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, subunit B'"
/protein_id="AAB99043.1"
/db_xref="GI:1591693"
/translation="MAIRKELVDVAFKEHGLIDHOIESNDFVNRILOKIILIDPVYIE
TETIGYKIKIGIKIVKCKPVKEADSGIRPITTEMRKIRIDLTYSPLXLMKPIIGEG
EDARBEIVYITGELPYMLGSKICHLYGKSRRELLIDJEDDEDFGFIITNGTEVL
ITQEDDILPNRILICEAKERSGKIVDVAVFSTRGFALCTYERHFDGLIATYPPGNG
QIDPVLIMALGATDNDKDIIESIDERFMEIVLITQETREHNTNSPEALERTGKR
VAGQAKDRLRAETVELCNLPLPHGVTKEDEPKIRIPLGIMARNALIELYPCYRGED
DKOHVYKRAKLAGDLMDCNLFYVAFSGIVYDKYQLEQOTLRNKTPSLQAAVSDLT
ERIKHMAATGTVGKTSVGSOLDIRSYLATNSQLRIVSPISRSQREARERLRGTH
WKLICSETPEGPCGVLNKAIMCKVTRREDSKVILELSEFINSV"
Alignment Scores:
Pred. No.: 162
Score: 41.00
Percent Similarity: 100.00%
Best Local Similarity: 72.73%
Query Match: 83.67%
Length: 10673
Matches: 8
Conservative: 3
Mismatches: 0
Indels: 0

DB:	1	Gaps:	0
US-09-823-649A-6 (1-11) x U67546 (1-10677)			
QY	1	LeuSerLySArgIleGlyLeuSerValSerIu 11	
Db	1320	CTCTCTAAAAAATGGATTTCTGTGCTCAT	1352
RESULT 7			
AX059549/c	47310 bp	DNA	linear
LOCUS			PAT 17-JAN-2001
DEFINITION	Sequence 282 from Patent WO0055325.		
ACCESSION	AX059549		
VERSION	AX059549.1		
KEYWORDS	GI:12311656		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	1 (bases 1 to 47310)		
AUTHORS	Preuss,D., Copenhaver,G. and Keith,K.		
TITLE	Plant chromosome compositions and methods		
JOURNAL	Patent: WO 0055325-A 282-21-SRP-2000;		
FEATURES	The University of Chicago (US)		
source	location/Qualifiers		
	1..47310		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
BASE COUNT	15441 a 8769 c 8265 g 14835 t		
ALIGN			
Alignment Scores:			
Pred. No.:	903	Length:	47310
Score:	41.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	90.00%	Mismatches:	0
Query Match:	83.67%	Indels:	0
DB:	6	Gaps:	0
US-09-823-649A-6 (1-11) x AX059549 (1-47310)			
OY	1	LeuSerLySArgIleGlyLeuSerValSer 10	
Db	43645	CTTCCAAACGATTGGCGTCCGGTCA	43616
RESULT 8			
FL14G16/c	63292 bp	DNA	linear
LOCUS			PLN 13-AUG-1999
DEFINITION	Arabidopsis thaliana BAC FL14G16.		
ACCESSION	AF147260		
VERSION	AF147260.1		
KEYWORDS	GI:4732165		
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE	1 (bases 1 to 63292)		
AUTHORS	Harmon,G., Wohldmann,P. and Lehnert,L.		
TITLE	The sequence of A. thaliana FL14G16		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 63292)		
AUTHORS	Washington University Genome Sequencing Center.		
TITLE	The A. thaliana Genome Sequencing Project		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 63292)		
AUTHORS	Waterston,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-APR-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, MO 63108, USA		
REFERENCE	4 (bases 1 to 63292)		
AUTHORS	Waterston,R.		

```

TITLE      Direct Submission
JOURNAL    Submitted (11-MAY-1999) Department of Genetics, Washington
REFERENCE  University, 4444 Forest Park Avenue, St. Louis, MO 63108, USA
AUTHORS    5 (bases 1 to 63292)
TITLE      Waterston,R.
JOURNAL    Direct Submission
COMMENT     Submitted (13-AUG-1999) Department of Genetics, Washington
           University, 4444 Forest Park Avenue, St. Louis, MO 63108, USA
           Submitted by:
           Genome Sequencing Center
           Department of Genetics, Washington University,
           St. Louis, MO 63108, USA
           e-mail: twilson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by
M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted
by M. Marra, Washu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

      NEIGHBORING COSMID INFORMATION:

The 3' clone is F28D6, 200 bp overlap. Actual start of this clone
is at base position 1 of F14G16; actual end is at 41649 of F28D6

NOTES:

Coding sequences below are predicted from computer analysis, using
the program GenefinderP. Green and L. Hillier, ms in preparation)

FEATURES
source
1..63292
   /organism="Arabidopsis thaliana"
   /cultivar="Columbia"
   /db_xref="taxon:3702"
   /chromosome="IV"
   /map="unknown"
   /clone="F14G16"
13..44
   /note="F14G16"
35..209
   /note="F14G16"
48..97
   /note="F14G16"
121..218
   /note="F14G16"
190..600
   /note="F14G16"
216..297
   /note="F14G16"
292..355
   /note="F14G16"
390..448
   /note="F14G16"
492..794
   /note="F14G16"
591..632
   /note="F14G16"
628..1027
   /note="F14G16"
987..1521

```

```

repeat_region /note="F14G16"
997.1578
repeat_region /note="F14G16"
1528.1684
repeat_region /note="F14G16"
1561.1647
repeat_region /note="F14G16"
1646.1907
repeat_region /note="F14G16"
1672.1814
repeat_region /note="F14G16"
1824.2012
repeat_region /note="F14G16"
1901.1977
repeat_region /note="F14G16"
1975.2479
repeat_region /note="F14G16"
1996.2315
repeat_region /note="F14G16"
2213.2530
repeat_region /note="F14G16"
2465.2565
repeat_region /note="F14G16"
2474.2791
repeat_region /note="F14G16"
2793.2826
repeat_region /note="F14G16"
2826.3601
repeat_region /note="F14G16"
3602.4665
repeat_region /note="F14G16"
5136.5243
repeat_region /note="F14G16"
5241.5443
repeat_region /note="F14G16"
5887.6013
repeat_region /note="F14G16"
6002.6401
repeat_region /note="F14G16"
7139.7282
repeat_region /note="F14G16"
7300.7710
repeat_region /note="F14G16"
7701.7747
repeat_region /note="F14G16"
7738.7837
repeat_region /note="F14G16"
7835.8136
repeat_region /note="F14G16"
8114.8270
repeat_region /note="F14G16"
10255.10398
repeat_region /note="F14G16"
10261.10435
repeat_region /note="F14G16"
10424.10854
repeat_region /note="F14G16"
10824.10870
repeat_region /note="F14G16"
10824.10865
repeat_region /note="F14G16"
10860.11259
repeat_region /note="F14G16"
10860.10959
repeat_region /note="F14G16"
10958.11259
repeat_region /note="F14G16"
11231.11812
repeat_region /note="F14G16"
11238.11394
repeat_region /note="F14G16"
11820.11906
repeat_region /note="F14G16"

```

```

repeat_region 11906.12167
/note="T32N15_del_retroTn"
repeat_region 12162.12238
/note="T32N15_del_retroTn"
repeat_region 12237.12741
/note="T32N15_del_retroTn"
misc_feature complement(12310..15735)
/note="pseudogene; similar to retrovirus-related
polyproteins"
repeat_region 12735.13052
/note="T32N15_del_retroTn"
repeat_region 13053.13086
/note="T32N15_del_retroTn"
repeat_region 13087.13862
/note="T32N15_del_retroTn"
repeat_region 13864.14927
/note="T32N15_del_retroTn"
repeat_region 14920.15027
/note="T32N15_del_retroTn"
repeat_region 15024.15226
/note="T32N15_del_retroTn"
repeat_region 15224.15350
/note="T32N15_del_retroTn"
repeat_region 15340.15739
/note="T32N15_del_retroTn"
gene /note="F14G16.2"
21183.22195
/note="F14G16.2"
CDS join(21183..21320,21398..21662,21795..21897,21973..22195)
/gene="F14G16.2"

```

```

Alignment Scores:
Pred. No.: 1.26e+03 Length: 63292
Score: 41.00 Matches: 9
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 90.00% Mismatches: 0
Query Match: 83.67% Indels: 0
DB: Gaps: 0

```

US-09-823-649A-6 (1-11) x F14G16 (1-63292)

Qy 1 LeuSerLysArgIleGlyLeuSerValSer 10

Db 43645 CTTTCAAACGATTTGGCGTCCGTGTA 43616

RESULT 9

AC113735/C

LOCUS AC113735

DEFINITION Rattus norvegicus clone CH230-20716. *** SEQUENCING IN PROGRESS

ACCESSION AC113735 120885 bp DNA linear HTG 17-JUL-2002

VERSION AC113735.4 GI:21450369

KEYWORDS HTG; HTGS; PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 120885)

AUTHORS

Munhy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Albrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T.,
 Baderia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Boucek,J., Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P.,
 Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferriguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Howett, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W., Loussened, H.,
 Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mel, G., Metker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okwuon, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Plickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y.,
 Rives, M., Rojals, A., Rojuckan, I., Rolfe, M., Ruiz, S., Severy, G.,
 Scherer, S., Scott, G., Shen, H., Shooshbari, N., Sisson, I.,
 Sodergren, E., Sotake, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Tellrod, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, Y., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
 Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and Gibbs, R.

 Direct Submission
 Unpublished
 2 (bases 1 to 120885)

 Direct Submission
 Submitted (05-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 120885)

 Direct Submission
 Submitted (17-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jun 18, 2002 this sequence version replaced gi:19525901.

 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GSDR
 Center clone name: CH230-20716

 Summary Statistics
 Sequencing vector: Plasmid:
 Chemistry: Dye-terminator Big Dye, 100% of reads
 Assembly program: Phrap, version 0.990329
 Consensus quality: 57067 bases at least Q40
 Consensus quality: 59738 bases at least Q30
 Consensus quality: 62357 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 55 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1059: contig of 1059 bp in length
 * 1060 1159: gap of unknown length
 * 1160 2297: contig of 1138 bp in length
 * 2298 2397: gap of unknown length
 * 2398 3499: contig of 1102 bp in length
 * 3500 3599: gap of unknown length
 * 3600 5048: contig of 1449 bp in length
 *
 * 5049 5148: gap of unknown length
 * 5149 6817: contig of 1669 bp in length
 * 6818 6917: gap of unknown length
 * 6918 7929: contig of 1012 bp in length
 * 7930 8029: gap of unknown length
 * 8030 9205: contig of 1176 bp in length
 * 9206 9305: gap of unknown length
 * 9306 10879: contig of 1574 bp in length
 * 10880 10979: gap of unknown length
 * 10980 12086: contig of 1107 bp in length
 * 12087 12186: gap of unknown length
 * 12187 13646: contig of 1460 bp in length
 * 13647 13746: gap of unknown length
 * 13747 14840: contig of 1094 bp in length
 * 14841 14940: gap of unknown length
 * 14941 16164: contig of 1224 bp in length
 * 16165 16265: gap of unknown length
 * 16265 17965: contig of 1701 bp in length
 * 17966 18065: gap of unknown length
 * 18066 19370: contig of 1305 bp in length
 * 19371 19470: gap of unknown length
 * 19471 20728: contig of 1258 bp in length
 * 20729 20828: gap of unknown length
 * 20829 21888: contig of 1160 bp in length
 * 21889 22088: gap of unknown length
 * 22089 23098: contig of 1010 bp in length
 * 23099 23198: gap of unknown length
 * 23199 24459: contig of 1261 bp in length
 * 24460 24590: gap of unknown length
 * 24591 26190: contig of 1631 bp in length
 * 26191 26290: gap of unknown length
 * 26291 27310: contig of 1020 bp in length
 * 27311 27410: gap of unknown length
 * 27411 28590: contig of 1180 bp in length
 * 28591 28690: gap of unknown length
 * 28691 29859: contig of 1269 bp in length
 * 29860 30059: gap of unknown length
 * 30060 31238: contig of 1179 bp in length
 * 31239 31338: gap of unknown length
 * 31339 32795: contig of 1457 bp in length
 * 32796 32895: gap of unknown length
 * 32896 34340: contig of 1445 bp in length
 * 34341 34440: gap of unknown length
 * 34441 35758: contig of 1318 bp in length
 * 35759 35858: gap of unknown length
 * 35859 37659: contig of 1801 bp in length
 * 37660 37759: gap of unknown length
 * 37760 38865: contig of 1106 bp in length
 * 38866 38965: gap of unknown length
 * 38966 40780: contig of 1815 bp in length
 * 40781 40880: gap of unknown length
 * 40881 42261: contig of 1381 bp in length
 * 42262 42361: gap of unknown length
 * 42362 44332: contig of 1971 bp in length
 * 44333 44432: gap of unknown length
 * 44433 46235: contig of 1803 bp in length
 * 46236 46335: gap of unknown length
 * 46336 47794: contig of 1455 bp in length
 * 47795 47894: gap of unknown length
 * 47895 49508: contig of 1614 bp in length
 * 49509 49608: gap of unknown length
 * 49609 51829: contig of 2221 bp in length
 * 51830 51929: gap of unknown length
 * 51930 53250: contig of 1321 bp in length
 * 53251 53350: gap of unknown length
 * 53351 54693: contig of 1343 bp in length
 * 54694 54793: gap of unknown length
 * 54794 56354: contig of 1561 bp in length
 * 56355 56454: gap of unknown length
 * 56455 58927: contig of 2473 bp in length
 * 58928 59027: gap of unknown length
 * 59028 61203: contig of 2176 bp in length
 * 61204 61303: gap of unknown length

Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Murinae;
Rattus.

1 (bases 1 to 185133)

Muzny,D.M., Adams,C., Adlo-otunola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Aye,J.R., Ayele,M., Banks,T.,
Barbarta,J.J., Benton,J., Bismae,K., Blankenburg,K., Bonani,D.,
Bouck,T.J., Bowse,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burck,P., Burkett,C., Burfell,K.L., Byrd,A.N.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Day-carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Eggard,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
Gabriel,A., Gao,J., Garcia,A., Garner,T., Gazda,N., Gill,R.,
Gorell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

* 6644 6743: gap of unknown length
* 6744 8148: contig of 1405 bp in length
* 8149 8248: gap of unknown length
* 8249 9756: contig of 1508 bp in length
* 9757 9856: gap of unknown length
* 9857 10897: contig of 1041 bp in length
* 10898 10997: gap of unknown length
* 10998 12769: contig of 1772 bp in length
* 12770 12869: gap of unknown length
* 12870 14106: contig of 1237 bp in length
* 14107 14206: gap of unknown length
* 14207 16029: contig of 1823 bp in length
* 16030 16129: gap of unknown length
* 16130 17629: contig of 1500 bp in length
* 17630 17729: gap of unknown length
* 17730 18891: contig of 1162 bp in length
* 18892 18991: gap of unknown length
* 18992 20718: contig of 1727 bp in length
* 20719 20818: gap of unknown length
* 20819 22300: contig of 1482 bp in length
* 22301 22400: gap of unknown length
* 22401 23490: contig of 1090 bp in length
* 23491 23590: gap of unknown length
* 23591 24737: contig of 1147 bp in length
* 24738 24837: gap of unknown length
* 24838 26116: contig of 1279 bp in length
* 26117 26216: gap of unknown length
* 26217 27298: contig of 1082 bp in length
* 27299 27398: gap of unknown length
* 27399 28433: contig of 1035 bp in length
* 28434 28533: gap of unknown length
* 28534 30486: contig of 1953 bp in length
* 30487 30586: gap of unknown length
* 30587 32725: contig of 2139 bp in length
* 32726 32825: gap of unknown length
* 32826 35521: contig of 2696 bp in length
* 35522 35621: gap of unknown length
* 35622 37201: contig of 1580 bp in length
* 37202 37301: gap of unknown length
* 37302 38964: contig of 1663 bp in length
* 38965 39064: gap of unknown length
* 39065 40735: contig of 1672 bp in length
* 40737 40836: gap of unknown length
* 40837 42720: contig of 1884 bp in length
* 42721 42820: gap of unknown length
* 42821 44727: contig of 1907 bp in length
* 44728 44827: gap of unknown length
* 44828 47162: contig of 2335 bp in length
* 47163 47262: gap of unknown length
* 47263 49017: contig of 1755 bp in length
* 49018 49117: gap of unknown length
* 49118 51193: contig of 2076 bp in length
* 51194 51293: gap of unknown length
* 51294 53244: contig of 1951 bp in length
* 53245 53344: gap of unknown length
* 53345 55072: contig of 1728 bp in length
* 55073 55172: gap of unknown length
* 55173 58133: contig of 2961 bp in length
* 58134 58233: gap of unknown length
* 58234 61216: contig of 2983 bp in length
* 61217 61316: gap of unknown length
* 61317 63283: contig of 1967 bp in length
* 63284 63383: gap of unknown length
* 63384 67659: contig of 4276 bp in length
* 67660 67759: gap of unknown length
* 67760 70442: contig of 2683 bp in length
* 70443 70542: gap of unknown length
* 70543 72498: contig of 1956 bp in length
* 72499 72598: gap of unknown length
* 72599 75291: contig of 2693 bp in length
* 75292 75391: gap of unknown length
* 75392 77963: contig of 2572 bp in length
* 77964 78063: gap of unknown length

```

```

* 78064 80993: contig of 2930 bp in length
* 80994 81093: gap of unknown length
* 81094 82875: contig of 1782 bp in length
* 82876 82975: gap of unknown length
* 82976 85626: contig of 2651 bp in length
* 85627 85727: gap of unknown length
* 85728 88361: contig of 2635 bp in length
* 88362 88461: gap of unknown length
* 88462 91705: contig of 3244 bp in length
* 91706 91805: gap of unknown length
* 91806 95029: contig of 3224 bp in length
* 95030 95129: gap of unknown length
* 95130 99071: contig of 3942 bp in length
* 99072 99171: gap of unknown length
* 99172 103377: contig of 4206 bp in length
* 103378 103478: gap of unknown length
* 103479 107321: contig of 3843 bp in length
* 107322 107420: gap of unknown length
* 107421 111282: contig of 3862 bp in length
* 111283 111382: gap of unknown length
* 111383 115735: contig of 4353 bp in length
* 115736 115835: gap of unknown length
* 115836 120535: contig of 4700 bp in length
* 120536 120636: gap of unknown length
* 120637 124813: contig of 4178 bp in length

```

Alignment Scores:

```

Pred. No.: 4.36e+03 Length: 185133
Score: 41.00 Matches: 8
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 72.73% Mismatches: 0
Query Match: 83.67% Indels: 0
DB: 2 Gaps: 0

```

US-09-823-649a-6 (1-11) x AC095167 (1-185133)

QY 1 Leuserlystrg1leglyLeuserValserclu 11

DB 133930 CTGCACAAAGAGTGGCTGACGTCTCGAA 133898

RESULT 11

AC099720 189300 bp DNA linear HTS 18-NOV-2001

LOCUS Mus musculus clone RP23-415H15, WORKING DRAFT SEQUENCE, 10

DEFINITION unoriented pieces.

ACCESSION AC099720

VERSION AC099720.1 GI:16974219

KEYWORDS HTGS_PHASE1, HTGS_DRAFT, HTGS_FULLTOP.

SOURCE Mus musculus.

ORGANISM Mus musculus.

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 189300)

JOURNAL Birren B., Linton L., Nusbaum C. and Lander E.

AUTHORS Mus musculus, clone RP23-415H15

REFERENCE 2 (bases 1 to 189300)

UNPUBLISHED

Birren B., Linton L., Nusbaum C., Lander E., Allen N.,

Anderson S., Baran N., Bastien V., Boguslavsky L., Bouckalter B.,

Brown A., Camarata J., Campiano A., Chang J., Chazaro B.,

Choquel Y., Colangelo M., Collins S., Collymore A., Cook A.,

Cooke P., DeRellano K., Dewar K., Diaz J.S., Dodge S., Fero S.,

Ferreira P., FitzHugh W., Gage D., Galagan J., Gardyna S.,

Grinde S., Gord S., Goyette M., Graham L., Grand-Pierre N.,
Hagos B., Heaford A., Horton L., Hulme W., Iliev I., Johnson R.,
Jones C., Kanalt A., Karatas A., Kelis C., Larocque K.,
Lamazares R., Landers T., Lehoczy J., Levine R., Liu G.,
Maclean C., Macdonald P., Major J., Marguis N., Matthews C.,
McCarthy M., McEwan P., McKernan K., McPeeters R., Melidze J.,
Meneus L., Mihova T., Mlenga V., Murphy J., Naylor J., Nguyen C.,
Norbu C., Norman C.H., O'Connor T., O'Donnell P., O'Neill D.,
Oliver C., Peterson K., Phunkhang P., Pierre N., Pollara V.,
Raymond C., Retta R., Rieback M., Riley R., Rise C., Rogov P.,

/note="F14G16"

```

FEATURES
source
1. 198176
   /organism="Arabidopsis thaliana"
   /variety="Columbia"
   /db_xref="taxon:3702"
   /chromosome="4"
13. 44
   /note="P14G16"
35. 209
   /note="P14G16"
repeat_unit
repeat_unit

```

[illegible]


```

* 15378 17818: contig of 2441 bp in length
* 17819 17818: gap of unknown length
* 17919 20608: contig of 2690 bp in length
* 20609 20708: gap of unknown length
* 20709 22732: contig of 2024 bp in length
* 22733 22832: gap of unknown length
* 22833 25082: contig of 2250 bp in length
* 25083 25182: gap of unknown length
* 25183 26421: contig of 1239 bp in length
* 26422 26521: gap of unknown length
* 26522 26826: contig of 3305 bp in length
* 26827 29926: gap of unknown length
* 29927 32988: contig of 3062 bp in length
* 32989 33088: gap of unknown length
* 33089 34562: contig of 1474 bp in length
* 34563 34662: gap of unknown length
* 34663 36581: contig of 1919 bp in length
* 36582 36681: gap of unknown length
* 36682 39244: contig of 2563 bp in length
* 39245 39345: gap of unknown length
* 42446 42545: contig of 3101 bp in length
* 42546 46046: gap of unknown length
* 46047 46146: gap of unknown length
* 46147 49114: contig of 2968 bp in length
* 49115 49214: gap of unknown length
* 49215 53838: contig of 4624 bp in length
* 53839 53938: gap of unknown length
* 53939 57266: contig of 3328 bp in length
* 57267 57367: gap of unknown length
* 57367 62101: contig of 4734 bp in length
* 62101 62200: gap of unknown length
* 62201 65616: contig of 3415 bp in length
* 65616 65715: gap of unknown length
* 65715 70020: contig of 4305 bp in length
* 70021 70120: gap of unknown length
* 70121 75631: contig of 5511 bp in length
* 75632 75731: gap of unknown length
* 75732 80827: contig of 5096 bp in length
* 80828 80927: gap of unknown length
* 80928 88473: contig of 7546 bp in length
* 88474 88573: gap of unknown length
* 88574 94842: contig of 6266 bp in length
* 94843 94942: gap of unknown length
* 94943 103669: contig of 8727 bp in length
* 103670 103770: gap of unknown length
* 103770 110999: contig of 7230 bp in length
* 111000 111099: gap of unknown length
* 111100 116739: contig of 5639 bp in length
* 116739 116838: gap of unknown length
* 116839 125092: contig of 8254 bp in length
* 125093 125192: gap of unknown length
* 125193 136630: contig of 11438 bp in length
* 136631 136730: gap of unknown length
* 136731 150018: contig of 13288 bp in length.

```

FEATURES

source 1.150018

Location/Qualifiers

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone="CH230-257F24"

BASE COUNT 41433 a 32641 c 31289 g 39633 t 5022 others

ORIGIN

Alignment Scores:

Pred. No.: 5.87e+03 Length: 150018

Score: 40.00 Matches: 8

Percent Similarity: 90.91% Conservative: 2

Best Local Similarity: 72.73% Mismatches: 1

Query Match: 81.63% Indels: 0

DB: 2 Gaps: 0

US-09-823-649a-6 (1-11) x AC114382 (1-150018)

Qy 1 LeuserTysArgTlleGlyLeuserValSerGlu 11

Db 30376 TTTAAGAGACAGATTGCAGCTTACAGTTTCGAG 30344

RESULT 14

AC107106

LOCUS

DEFINITION

AC107106 170083 bp DNA linear HTG 17-JUL-2002

Rattus norvegicus clone CH230-155W2, *** SEQUENCING IN PROGRESS

***, 61 unordered pieces.

ACCESSION

AC107106

VERSION

AC107106.3 GI:21743917

KEYWORDS

HTG; HTGS_PHASE1.

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 170083)

Muzny,D.M., Adams,C., Adio-Otunola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,

Barbieri,J., Benton,J., Bimaye,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

DeVila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Correll,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,

HomsI,F., Howard,S., Huber,J., Hult,K.S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,

Kratovic,J., Kurshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,

Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhney,E., McLeod,M.P., Medor,M., Mel,G., Melzer,M.,

Miner,G., Miner,Z., Mitchell,T., Nohshat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokkenwo,S., Ogum,M., Okunom,G.,

Oranunye,N., Oviedo,K., Pace,A., Payton,B., Percy,J., Perez,L.,

Peters,L., Pickens,K., Primus,E., Pul,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Saverly,G.,

Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,

Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Usmani,K., Vasquez,L., Vera,Y., Villalob,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,

Williams,G., Williamson,A., Wlezyk,R., Wooten,S., Worley,K.,

Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G., and Gibbs,R.

TITLE

Direct Submission

REFERENCE

2 (bases 1 to 170083)

AUTHORS

Worley,K.C.

TITLE

Submitted (16-JAN-2002) Human Genome Sequencing Center, Department

REFERENCE

3 (bases 1 to 170083)

AUTHORS

Worley,K.C.

TITLE

Submitted (17-JUL-2002) Human Genome Sequencing Center, Department

COMMENT

On Jul 13, 2002 this sequence version replaced g1:18701531.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNIF
Center clone name: CH230-155M2
----- Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 111369 bases at least Q40
Consensus quality: 116447 bases at least Q30
Consensus quality: 121122 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1039: contig of 1039 bp in length
* 1040 1139: gap of unknown length
* 1140 2259: contig of 1120 bp in length
* 2260 2359: gap of unknown length
* 2360 3500: contig of 1141 bp in length
* 3501 3600: gap of unknown length
* 3601 4883: contig of 1283 bp in length
* 4884 4983: gap of unknown length
* 4984 6335: contig of 1352 bp in length
* 6336 6435: gap of unknown length
* 6436 7741: contig of 1306 bp in length
* 7742 7841: gap of unknown length
* 7842 9151: contig of 1310 bp in length
* 9152 9251: gap of unknown length
* 9252 10753: contig of 1502 bp in length
* 10754 10853: gap of unknown length
* 10854 12543: contig of 1690 bp in length
* 12544 12643: gap of unknown length
* 12644 13865: contig of 1222 bp in length
* 13866 13965: gap of unknown length
* 13966 15202: contig of 1237 bp in length
* 15203 15302: gap of unknown length
* 15303 16317: contig of 1015 bp in length
* 16318 16417: gap of unknown length
* 16418 17872: contig of 1455 bp in length
* 17873 17972: gap of unknown length
* 17973 19199: contig of 1227 bp in length
* 19200 19299: gap of unknown length
* 19300 20563: contig of 1264 bp in length
* 20564 20663: gap of unknown length
* 20664 22126: contig of 1463 bp in length
* 22127 22226: gap of unknown length
* 22227 23739: contig of 1513 bp in length
* 23740 23839: gap of unknown length
* 23840 25334: contig of 1495 bp in length
* 25335 25434: gap of unknown length
* 25435 27296: contig of 1862 bp in length
* 27297 27396: gap of unknown length
* 27397 28963: contig of 1567 bp in length
* 28964 29063: gap of unknown length
* 29064 30622: contig of 1559 bp in length
* 30623 30722: gap of unknown length
* 30723 31939: contig of 1217 bp in length
* 31940 32039: gap of unknown length
* 32040 34250: contig of 2211 bp in length
* 34251 34351: gap of unknown length
* 34351 36580: contig of 2230 bp in length

```

```

* 36581 36680: gap of unknown length
* 36681 38029: contig of 1349 bp in length
* 38030 38129: gap of unknown length
* 38130 40235: contig of 2106 bp in length
* 40236 40335: gap of unknown length
* 40336 42225: contig of 1890 bp in length
* 42226 42325: gap of unknown length
* 42326 44117: contig of 1792 bp in length
* 44118 44217: gap of unknown length
* 44218 45978: contig of 1761 bp in length
* 45979 46078: gap of unknown length
* 46079 47659: contig of 1581 bp in length
* 47660 47759: gap of unknown length
* 47760 49260: contig of 1501 bp in length
* 49261 49360: gap of unknown length
* 49361 51125: contig of 1765 bp in length
* 51126 51225: gap of unknown length
* 51226 54076: contig of 2851 bp in length
* 54077 54176: gap of unknown length
* 54177 55760: contig of 1584 bp in length
* 55761 55860: gap of unknown length
* 55861 58528: contig of 2668 bp in length
* 58529 58628: gap of unknown length
* 58629 61885: contig of 3257 bp in length
* 61886 61985: gap of unknown length
* 61986 65195: contig of 3210 bp in length
* 65196 65295: gap of unknown length
* 65296 67635: contig of 2340 bp in length
* 67636 67735: gap of unknown length
* 67736 70102: contig of 2367 bp in length
* 70103 70202: gap of unknown length
* 70203 73728: contig of 3526 bp in length
* 73729 73828: gap of unknown length
* 73829 76882: contig of 3054 bp in length
* 76883 76982: gap of unknown length
* 76983 80139: contig of 3157 bp in length
* 80140 80239: gap of unknown length
* 80240 82873: contig of 2634 bp in length
* 82874 82973: gap of unknown length
* 82974 85025: contig of 2052 bp in length
* 85026 85125: gap of unknown length
* 85126 89485: contig of 4370 bp in length
* 89486 89585: gap of unknown length
* 89586 92436: contig of 2741 bp in length
* 92437 95301: contig of 2865 bp in length
* 95302 95401: gap of unknown length
* 95402 100512: contig of 5111 bp in length
* 100513 100612: gap of unknown length
* 100613 104557: contig of 2945 bp in length
* 104558 104657: gap of unknown length
* 104658 107805: contig of 4148 bp in length
* 107806 107905: gap of unknown length
* 107906 110962: contig of 3057 bp in length
* 110963 111062: gap of unknown length
* 111063 11462: contig of 3400 bp in length
* 11463 114662: gap of unknown length
* 114663 117995: contig of 3433 bp in length
* 117996 118095: gap of unknown length
* 118096 121626: contig of 3531 bp in length

```

Alignment Scores:

```

Pred. No.: 6.78e+03 Length: 170083
Score: 40.00 Matches: 7
Percent Similarity: 100.00% Conservative: 3
Best local Similarity: 70.00% Mismatches: 0
Query Match: 81.63% Indels: 0
DB: 2 Gaps: 0

```

US-09-823-649A-6 (1-11) x AC107106 (1-170083)

Oy 2 Ser1ysArg11Iec1y1LeuSerValSerGlu 11
 |||||.....

DB 128006 TCTAGAGACTAGCTTGACATTACGAA 128035

RESULT 15 AC094431/c

LOCUS AC094431 180674 bp DNA linear HTG 10-JUL-2002
DEFINITION Rattus norvegicus clone CH230-3012, *** SEQUENCING IN PROGRESS ***,
65 unordered pieces.

ACCESSION AC094431
VERSION AC094431.3 GI:21716471
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE AUTHORS

1 (bases 1 to 180674)

Alzryu,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Mishra,S.L., Amarasinghe,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buday,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Duttwalte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,D., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivel,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Koryah,J., Kovar,C.,
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,
Lc Jdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Mashwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Melker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,R., Nwokenwo,S., Oguh,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Kojunokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonalke,A., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,O.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlezyk,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
TITLE
AUTHORS
JOURNAL

Unpublished
2 (bases 1 to 180674)
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
JOURNAL
COMMENT
On Jul 9, 2002 this sequence version replaced gi:17941157.
Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GAOY
Center clone name: CH230-3012
Summary Statistics

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye 100k of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124332 bases at least Q40
Consensus quality: 129469 bases at least Q30
Consensus quality: 134129 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently
consists of 65 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1	1363:	contig of 1363 bp in length
1364	1463:	gap of unknown length
1464	2556:	contig of 1093 bp in length
2557	2656:	gap of unknown length
2657	3761:	contig of 1105 bp in length
3762	3861:	gap of unknown length
3862	5074:	contig of 1213 bp in length
5075	5174:	gap of unknown length
5175	6113:	contig of 1239 bp in length
6114	6513:	gap of unknown length
6514	7972:	contig of 1459 bp in length
7973	8072:	gap of unknown length
8073	9421:	contig of 1349 bp in length
9422	9521:	gap of unknown length
9522	10575:	contig of 1034 bp in length
10576	10675:	gap of unknown length
10676	12022:	contig of 1347 bp in length
12023	12122:	gap of unknown length
12123	13331:	contig of 1109 bp in length
13332	13331:	gap of unknown length
13333	14441:	contig of 1110 bp in length
14442	14541:	gap of unknown length
14542	16334:	contig of 1793 bp in length
16335	16334:	gap of unknown length
16335	17651:	contig of 1217 bp in length
17652	17751:	gap of unknown length
17752	19601:	contig of 1850 bp in length
19602	19701:	gap of unknown length
19702	21524:	contig of 1823 bp in length
21525	21624:	gap of unknown length
21625	23157:	contig of 1533 bp in length
23158	23257:	gap of unknown length
23258	24316:	contig of 1059 bp in length
24317	24416:	gap of unknown length
24417	26018:	contig of 1602 bp in length
26019	26118:	gap of unknown length
26119	28044:	contig of 1926 bp in length
28045	28144:	gap of unknown length
28145	30555:	contig of 2411 bp in length
30556	30655:	gap of unknown length
30656	33487:	contig of 2832 bp in length
33488	33587:	gap of unknown length
33588	34946:	contig of 1359 bp in length
34947	35046:	gap of unknown length
35047	37112:	contig of 2066 bp in length
37113	37212:	gap of unknown length
37213	38658:	contig of 1646 bp in length
38659	38958:	gap of unknown length
38959	40723:	contig of 1765 bp in length

Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 16 Gaps: 0

US-09-823-649A-7 (1-11) x AAT04800 (1-2631)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArqLYsGlu 11
|||||
DB 2161 CTGGCGCAAACTTGACATFACGGCGCAAGAA 2193

Search completed: January 15, 2003, 12:56:57
Job time : 136.143 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 16, 2003, 08:03:43 : Search time 1525.14 Seconds
(without alignments)
209.902 Million cell updates/sec

Title: US-09-823-649a-7
Perfect score: 49
Sequence: 1 LAQNINXRKE 11

Scoring table: BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 1455102878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters: --
-MODEL-frame+ p2n.model -DEV-yth
-Q/cg22_1/usptc_sptc/us0982649/runat_14012003_151002_29139/app.query.fasta_1.1393
-DB-genemb1 -QFMT-fastlap -SUFFIX-rge -MINMATCH-0.1 -DOOPCL-0 -DOOEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST-45
-DOCALIGN-200 -THR-SCORE-pcl -THR-MAX-100 -THR-MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pco -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-us0982649 @CGN 1.1 3378 @runat 14012003 151002 29139 -NCPU-6 -ICPU-3
-NO_XLPEX -NO_MAP -LARGOQUERY -NRG-SCORES-0 -NAIT -LONGLOG -DEV-TIMMOUT=120
-MARK-TIMEOUT=30 -THREDS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : GenEmb1:*
1: gb.ba:*
2: gb.btg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg_hum:*
31: em.htg_inv:*
32: em.htg_other:*
33: em.htg_mus:*
34: em.htg_pin:*
35: em.htg_rod:*
36: em.htg_mam:*
37: em.htg_vrt:*
38: em.sy:*
39: em.htgo_hum:*
40: em.htgo_mus:*
41: em.htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	48	98.0	890 6	AX002399
2	48	98.0	1142 6	AX002398
3	48	98.0	1379 6	AX002397
4	48	98.0	1625 6	AX002396
5	48	98.0	1764 6	AR053713
6	48	98.0	1764 6	AR122765
7	48	98.0	1764 6	AR122766
8	48	98.0	1770 6	AR053714
9	48	98.0	1779 6	AR008369
10	48	98.0	1779 6	E05777
11	48	98.0	1779 6	I13466
12	48	98.0	1880 6	AX002395
13	48	98.0	2628 6	E05854
14	48	98.0	2631 6	BS033536
15	48	98.0	2631 6	AR008368
16	48	98.0	2631 6	AR043604
17	48	98.0	2631 6	E05778
18	48	98.0	2631 6	E13101
19	48	98.0	2631 6	I13465
20	48	98.0	2761 1	BACPOL
21	48	98.0	2814 1	BS093028
22	48	98.0	2969 1	BS023149
23	48	98.0	2969 6	AX002388
24	48	98.0	3246 6	E05853
25	48	98.0	3246 6	AR008367
26	48	98.0	4252 6	E05902
27	48	98.0	4252 6	I13464
28	48	98.0	4317 6	AK051777
29	48	98.0	3329 1	HACPOLYTC
30	41	83.7	138036 2	AP004670
31	41	83.7	177664 2	AP004620
32	40	81.6	47445 2	AC017438
33	40	81.6	185722 3	AC009341
34	40	81.6	261000 3	AE003654
35	40	81.6	301550 1	AP003192
36	39	79.6	2083 6	AX416375
37	39	79.6	2545 6	AX433109
38	39	79.6	2631 6	AR000463
39	39	79.6	2631 6	E23276
40	39	79.6	3255 6	AR000464
41	39	79.6	3255 6	E23277
42	39	79.6	3831 6	AR000465
43	39	79.6	3831 6	E23278
44	39	79.6	36052 3	I14429
45	39	79.6	67845 2	AC016315

ALIGNMENTS

RESULT 1

AX002399
LOCUS AX002399 890 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 12 from Patent EP0875576.
ACCESSION AX002399
VERSION AX002399.1 GI:7241992
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 890)
AUTHORS Tan, T.C. and Phang, S.M.
TITLE Bacillus stearothermophilus DNA polymerase I (Klenow) clones including those with reduced 3'-to-5' exonuclease activity
JOURNAL Patent: EP 0875576-A 12 04-NOV-1998;
UNIV SINGAPORE (SG)
FEATURES
source location/Qualifiers
1..890
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 227 a 228 c 257 g 178 t
ORIGIN
Alignment Scores:
Pred. No.: 0.138 Length: 890
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: Gaps: 0
US-09-823-649A-7 (1-11) x AX002399 (1-890)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArglySGlu 11
Db 325 CTGGCGCAAACTTGACATTTACGCGCAAGAA 357
RESULT 2
LOCUS AX002398 1142 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 11 from Patent EP0875576.
ACCESSION AX002398
VERSION AX002398.1 GI:7241991
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1142)
AUTHORS Tan, T.C. and Phang, S.M.
TITLE Bacillus stearothermophilus DNA polymerase I (Klenow) clones including those with reduced 3'-to-5' exonuclease activity
JOURNAL Patent: EP 0875576-A 11 04-NOV-1998;
UNIV SINGAPORE (SG)
FEATURES
source location/Qualifiers
1..1142
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 299 a 292 c 321 g 230 t
ORIGIN
Alignment Scores:
Pred. No.: 0.182 Length: 1142
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: Gaps: 0
US-09-823-649A-7 (1-11) x AX002398 (1-1142)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArglySGlu 11
Db 577 CTGGCGCAAACTTGACATTTACGCGCAAGAA 609

RESULT 3
LOCUS AX002397 1379 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 10 from Patent EP0875576.
ACCESSION AX002397
VERSION AX002397.1 GI:7241990
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1379)
AUTHORS Tan, T.C. and Phang, S.M.
TITLE Bacillus stearothermophilus DNA polymerase I (Klenow) clones including those with reduced 3'-to-5' exonuclease activity
JOURNAL Patent: EP 0875576-A 10 04-NOV-1998;
UNIV SINGAPORE (SG)
FEATURES
source location/Qualifiers
1..1379
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 355 a 351 c 401 g 272 t
ORIGIN
Alignment Scores:
Pred. No.: 0.224 Length: 1379
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: Gaps: 0
US-09-823-649A-7 (1-11) x AX002397 (1-1379)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArglySGlu 11
Db 814 CTGGCGCAAACTTGACATTTACGCGCAAGAA 846
RESULT 4
LOCUS AX002396 1625 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 9 from Patent EP0875576.
ACCESSION AX002396
VERSION AX002396.1 GI:7241989
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1625)
AUTHORS Tan, T.C. and Phang, S.M.
TITLE Bacillus stearothermophilus DNA polymerase I (Klenow) clones including those with reduced 3'-to-5' exonuclease activity
JOURNAL Patent: EP 0875576-A 9 04-NOV-1998;
UNIV SINGAPORE (SG)
FEATURES
source location/Qualifiers
1..1625
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 402 a 413 c 492 g 318 t
ORIGIN
Alignment Scores:
Pred. No.: 0.269 Length: 1625
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: Gaps: 0
US-09-823-649A-7 (1-11) x AX002396 (1-1625)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArglySGlu 11
Db 1060 CTGGCGCAAACTTGACATTTACGCGCAAGAA 1092

RESULT 5
LOCUS AR053713 1764 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5834253.
ACCESSION AR053713
VERSION AR053713.1 GI:5978575
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1764)
AUTHORS Hong, G., Fan., Huang, W., -h. and Zhai, F. deceased.
TITLE Bacillus stearothermophilus DNA polymerase with proof-reading 3'-5'
JOURNAL Patent: US 5834253-A 1 10-NOV-1998;
FEATURES
source location/Qualifiers
1..1764
/organism="unknown"
BASE COUNT 477 a 415 c 486 g 386 t
ORIGIN
Alignment Scores:
Pred. No.: 0.294 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
Gaps: 0
DB: 6
US-09-823-649a-7 (1-11) x AR053713 (1-1764)
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
DB 1294 TTGGCGCAAACTTGACATTACCGCAAGAA 1326
RESULT 6
LOCUS AR122765 1764 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6165765.
ACCESSION AR122765
VERSION AR122765.1 GI:14107082
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1764)
AUTHORS Hong, G. and Huang, W., -h.
TITLE DNA polymerase having ability to reduce innate selective
discrimination against fluorescent dye-labeled dideoxynucleotides
JOURNAL Patent: US 6165765-A 1 26-DEC-2000;
FEATURES
source location/Qualifiers
1..1764
/organism="unknown"
BASE COUNT 477 a 415 c 486 g 386 t
ORIGIN
Alignment Scores:
Pred. No.: 0.294 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
Gaps: 0
DB: 6
US-09-823-649a-7 (1-11) x AR122765 (1-1764)
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
DB 1294 TTGGCGCAAACTTGACATTACCGCAAGAA 1326
RESULT 7
LOCUS AR122766 1764 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 5834253.
ACCESSION AR122766
VERSION AR122766.1 GI:14107083
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1764)
AUTHORS Hong, G. and Huang, W., -h.
TITLE DNA polymerase having ability to reduce innate selective
discrimination against fluorescent dye-labeled dideoxynucleotides
JOURNAL Patent: US 5834253-A 3 26-DEC-2000;
FEATURES
source location/Qualifiers
1..1764
/organism="unknown"
BASE COUNT 476 a 418 c 485 g 385 t
ORIGIN
Alignment Scores:
Pred. No.: 0.294 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
Gaps: 0
DB: 6
US-09-823-649a-7 (1-11) x AR122766 (1-1764)
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
DB 1297 TTGGCGCAAACTTGACATTACCGCAAGAA 1329
RESULT 8
LOCUS AR053714 1770 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5834253.
ACCESSION AR053714
VERSION AR053714.1 GI:5978576
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1770)
AUTHORS Hong, G., Fan., Huang, W., -h. and Zhai, F. deceased.
TITLE Bacillus stearothermophilus DNA polymerase with proof-reading 3'-5'
JOURNAL Patent: US 5834253-A 3 10-NOV-1998;
FEATURES
source location/Qualifiers
1..1770
/organism="unknown"
BASE COUNT 481 a 415 c 487 g 387 t
ORIGIN
Alignment Scores:
Pred. No.: 0.295 Length: 1770
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
Gaps: 0
DB: 6
US-09-823-649a-7 (1-11) x AR053714 (1-1770)
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
DB 1300 TTGGCGCAAACTTGACATTACCGCAAGAA 1332
RESULT 9
LOCUS AR008369 1779 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 16 from patent US 5753482.
ACCESSION AR008369

LOCUS AR122766 1764 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 5834253.
ACCESSION AR122766
VERSION AR122766.1 GI:14107083
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1764)
AUTHORS Hong, G. and Huang, W., -h.
TITLE DNA polymerase having ability to reduce innate selective
discrimination against fluorescent dye-labeled dideoxynucleotides
JOURNAL Patent: US 5834253-A 3 26-DEC-2000;
FEATURES
source location/Qualifiers
1..1764
/organism="unknown"
BASE COUNT 476 a 418 c 485 g 385 t
ORIGIN
Alignment Scores:
Pred. No.: 0.294 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
Gaps: 0
DB: 6
US-09-823-649a-7 (1-11) x AR122766 (1-1764)
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
DB 1297 TTGGCGCAAACTTGACATTACCGCAAGAA 1329
RESULT 8
LOCUS AR053714 1770 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5834253.
ACCESSION AR053714
VERSION AR053714.1 GI:5978576
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1770)
AUTHORS Hong, G., Fan., Huang, W., -h. and Zhai, F. deceased.
TITLE Bacillus stearothermophilus DNA polymerase with proof-reading 3'-5'
JOURNAL Patent: US 5834253-A 3 10-NOV-1998;
FEATURES
source location/Qualifiers
1..1770
/organism="unknown"
BASE COUNT 481 a 415 c 487 g 387 t
ORIGIN
Alignment Scores:
Pred. No.: 0.295 Length: 1770
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
Gaps: 0
DB: 6
US-09-823-649a-7 (1-11) x AR053714 (1-1770)
OY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
DB 1300 TTGGCGCAAACTTGACATTACCGCAAGAA 1332
RESULT 9
LOCUS AR008369 1779 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 16 from patent US 5753482.
ACCESSION AR008369

VERSION AR008369.1 GI:3967478
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE Unclassified.
 AUTHORS 1 (bases 1 to 1779)
 TITLE Ishino,Y., Uemori,T., Fujita,K. and Kato,I.
 JOURNAL Method for cloning of a gene for pol I type DNA polymerase
 FEATURES Patent: US 5753482-A 16 19-MAY-1998;
 SOURCE Location/Qualifiers
 1..1779
 /organism="unknown"
 BASE COUNT 446 a 425 c 529 g 379 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.297 Length: 1779
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 DB: 6 Gaps: 0
 US-09-823-649a-7 (1-11) x AR008369 (1-1779)
 QY 1 LeuAlaGlnAsnLeuAsnIle***ArgIysGlu 11
 Db 1312 TTGGCGCAAACTTAATATTTCACGCAAGAG 1344
 RESULT 10
 LOCUS E05777 1779 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA encoding a DNA polymerase.
 E05777
 E05777.1 GI:2173964
 JP 1993284971-A/1.
 KEYWORDS Bacillus caldotenax.
 SOURCE Bacillus caldotenax.
 ORGANISM Bacillus caldotenax.
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 REFERENCE 1 (bases 1 to 1779)
 AUTHORS Ishino,Y., Uemori,T., Fujita,Y. and Kato,I.
 TITLE DNA POLYMERASE GENE
 JOURNAL Patent: JP 1993284971-A 1 02-NOV-1993;
 TAKARA SHUZO CO LTD
 OS Bacillus caldotenax
 FM JP 1993284971-A/1
 PD 02-NOV-1993
 PF 06-APR-1992 JP 1992112400
 PI ISHINO YOSHIZUMI, UEMORI TAKASHI, FUJITA YOSHIYO, PI KATOU
 IKUNOSHIN
 PC C12N15/54,C12N9/12,C12N15/54,C12RL:07);
 CC strandedness: Double;
 CC topology: Linear;
 CC *source: strain=YT-G(DSM406);
 CC *source: clone=pu1205;
 FH key Location/Qualifiers
 FH mat_peptide 1..1779
 FT Location/Qualifiers
 1..1779
 /organism="Bacillus caldotenax"
 /db_xref="taxon:1395"
 BASE COUNT 446 a 425 c 529 g 379 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.297 Length: 1779
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 DB: 6 Gaps: 0

US-09-823-649a-7 (1-11) x E05777 (1-1779)
 QY 1 LeuAlaGlnAsnLeuAsnIle***ArgIysGlu 11
 Db 1312 TTGGCGCAAACTTAATATTTCACGCAAGAG 1344
 RESULT 11
 LOCUS I13466 1779 bp DNA linear PAT 26-JUL-1995
 DEFINITION Sequence 16 from patent US 5436326.
 I13466
 I13466.1 GI:910807
 ACCESSION 113466
 VERSION 113466.1 GI:910807
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1779)
 AUTHORS Ishino,Y., Uemori,T., Fujita,K. and Kato,I.
 TITLE Method for cloning of a gene for pol I type DNA polymerase
 JOURNAL Patent: US 5436326-A 16 25-JUL-1995;
 FEATURES Location/Qualifiers
 1..1779
 /organism="unknown"
 BASE COUNT 446 a 425 c 529 g 379 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.297 Length: 1779
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 DB: 6 Gaps: 0
 US-09-823-649a-7 (1-11) x AX002395 (1-1880)
 QY 1 LeuAlaGlnAsnLeuAsnIle***ArgIysGlu 11
 Db 1312 TTGGCGCAAACTTAATATTTCACGCAAGAG 1344
 RESULT 12
 LOCUS AX002395 1880 bp DNA linear PAT 10-MAR-2000
 DEFINITION Sequence 8 from Patent EP0875576.
 AX002395
 AX002395.1 GI:7241988
 ACCESSION AX002395
 VERSION AX002395.1 GI:7241988
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1880)
 AUTHORS Tan,T.C. and Phang,S.M.
 TITLE Bacillus stearothermophilus DNA polymerase I (Klenow) clones
 JOURNAL including those with reduced 3'-to-5' exonuclease activity
 Patent: EP 0875576-A 8 04-NOV-1998;
 UNIV SINGAPORE (SG)
 FEATURES Location/Qualifiers
 1..1880
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 458 a 486 c 573 g 363 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.316 Length: 1880
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 DB: 6 Gaps: 0

Qy 1 LeuA1aGlnAsnLeuAsnIle***ArgLysGlu 11
 |||||
 Db 1315 CTGGCCCAAACTTACACATTACGCCAAGAA 1347

RESULT 13
 E05854
 LOCUS 2628 bp DNA linear PAT 29-SEP-1997
 DEFINITION DNA polymerase gene.
 ACCESSION E05854
 VERSION E05854.1 GI:2174041
 KEYWORDS JP 199304964-A/4.
 SOURCE Unidentified.
 ORGANISM Geobacillus stearothermophilus
 Bacteria; Firmicutes; Bacillales; Geobacillus.
 REFERENCE 1 (bases 1 to 2628)
 AUTHORS Ishino, Y., Demori, T., Fujita, Y. and Katou, I.
 TITLE DNA POLYMERASE GENE
 JOURNAL Patent: JP 199304964-A 4 19-NOV-1993;
 TAKARA SHUZO CO LTD
 COMMENT PN JP 199304964-A/4
 PD 19-NOV-1993
 PF 27-APR-1992 JP 1992131400
 PI ISHINO YOSHIZUMI, DEMORI TAKASHI, FUJITA YOSHIYO, PI KATOU IKUNOSHIN
 PC C12N15/54, C12N9/12, (C12N15/54, C12N1:07);
 CC strandedness: Double;
 CC topology: linear;
 CC Feature is identified by experimental;
 FH key Location/Qualifiers
 FT CDS 1..2628
 /product="DNA polymerase"
 FEATURES
 source 1..2628
 Location/Qualifiers
 1..2628
 /organism="Geobacillus stearothermophilus"
 /db_xref="taxon:1422"

BASE COUNT 668 a 636 c 771 g 553 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.458 Length: 2628
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 Gaps: 0

US-09-823-649a-7 (1-11) x E05854 (1-2628)

Qy 1 LeuA1aGlnAsnLeuAsnIle***ArgLysGlu 11
 |||||
 Db 2161 TTGGCCCAAACTTAAATATTTCACGCAAGAG 2193

RESULT 14
 BSU33536
 LOCUS 2631 bp DNA linear BCT 21-APR-1997
 DEFINITION Geobacillus stearothermophilus DNA polymerase I (polA) gene, complete.
 cds.
 ACCESSION U33536
 VERSION U33536.1 GI:1205983
 KEYWORDS
 SOURCE
 ORGANISM Geobacillus stearothermophilus.
 Bacteria; Firmicutes; Bacillales; Geobacillus.
 REFERENCE 1 (bases 1 to 2631)
 AUTHORS Aliotta, J.M., Pelletier, J.J., Ware, J.L., Moran, L.S., Benner, J.S. and Kong, H.
 TITLE The most stable Bst DNA polymerase I lacks a 3'-->5' proofreading exonuclease activity
 JOURNAL Genet. Anal. 12 (5-6), 185-195 (1996)
 MEDLINE 96303305
 PUBMED 8740835

REFERENCE 2 (bases 1 to 2631)
 AUTHORS Kong, H.
 TITLE Direct Submission
 JOURNAL Submitted (04-MAY-1995) Huimin Kong, Research Department, New England Biolabs, 32 Tozer Road, Beverly, MA 01915, USA

FEATURES
 source 1..2631
 Location/Qualifiers
 1..2631
 /organism="Geobacillus stearothermophilus"
 /db_xref="taxon:1422"
 1..2631
 /gene="polA"
 1..2631
 /gene="polA"
 /codon_start=1
 /transl_table=11
 /product="DNA polymerase I"
 /protein_id="AAB52611.1"
 /db_xref="GI:1205984"
 /translation="MKKKIYLLDGNSSVYAPAPFLPLHNDKGLHTNAAVYGFMMMLNK
 IIAEOPPHILAAFPACKITFRHETPOEYKGGQCPPLSPHLSRPHLLBELIKAYRI
 YELDHYEADITIGTLAARBOGEFVYKISGPDQLQASRYVYDITKKGITDIPY
 TPETVREKGLTPBOIVDKGLMDKSDNTPGVIGETAVKLLKQGEVENVASI
 DVEYGEKLEKNRQHDLDLALLSKQASICDAVELEDDIVYEGDREYVIALFEL
 GPQSELEKMAAPAAAGCEKPLEMEFAIVADITTEMLADKAAALVYVMEENVHAPIVG
 IALVNEGRFPRPFTALADSOFLAMLADETKKSKSFARAVVALMKKCIERGVAF
 DILAAVLLNPADODAGDIAAFAKKOYBAVSPDEAYGKGRSLPDEOTIAELVAK
 AAALIMLEDPENDDLRNNEODQLITKEOPLAIIAEMETGVYVDRKRLBOMGSLA
 EQLRAIQRRIYELACQERNINSPKQIVITFENLOIPVTKKRTYISADVLEKAP
 HHETVEPILHTRKQLOSTYIEGLVAVPDTGKVHTFENALQTERLSSEPNIQ
 NPIRLIEGKRIQAVVSPEDMDLIFAADYSQLELVLAHIAADDNLIAEORNDIH
 TKTAMDIFHVSEEVYANMKRQAKAVNFIQISDYGIAQMLNITRKAARFIERF
 ASFGVQKYMENIVOEAKQKGYVTLLHRRRYLDTNSFNENRFAFRPAMNPLOG
 SAADIIRKAMIDILARKEBQLQRLLLQYHDELILEAKKEIENLELVEVMEQAV
 TLRVPLKVDYHNGPTWYDAK"

BASE COUNT 721 a 616 c 716 g 578 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.458 Length: 2631
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 Gaps: 0

US-09-823-649a-7 (1-11) x BSU33536 (1-2631)

Qy 1 LeuA1aGlnAsnLeuAsnIle***ArgLysGlu 11
 |||||
 Db 2161 TTGGCCCAAACTTAAACATTTACGCCAAGAA 2193

RESULT 15
 AR008368
 LOCUS 2631 bp DNA linear PAT 04-DEC-1998
 DEFINITION Sequence 13 from patent US 5753482.
 ACCESSION AR008368
 VERSION AR008368.1 GI:3967477
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2631)
 AUTHORS Ishino, Y., Demori, T., Fujita, K. and Kato, I.
 TITLE Method for cloning of a gene for pol I type DNA polymerase
 JOURNAL Patent: US 5753482-A 13 19-MAY-1998;
 FEATURES
 source 1..2631
 Location/Qualifiers
 1..2631
 /organism="unknown"
 BASE COUNT 667 a 638 c 774 g 552 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.458 Length: 2631

Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 6 Gaps: 0
US-09-823-649a-7 (1-11) x AR008368 (1-2631)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
|||||
DB 2164 TTGGCGCAAACTTAAATATTCACGCAAGAG 2196
Search completed: January 16, 2003, 11:05:00
Job time : 1534.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 10:29:39 : Search time 25.5714 Seconds
(without alignments)
88.635 Million cell updates/sec

Title: US-09-823-649a-7
Perfect score: 49
Sequence: 1 LAONLIXRKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	98.0	876	2 Q45458	Q45458 bacillus st
2	48	98.0	877	2 Q24675	Q24675 bacillus st
3	40	81.6	866	16 Q8XW9	Q8XW9 clostridium
4	39	79.6	875	16 Q92BFO	Q92BFO listeria in
5	39	79.6	875	16 Q8T6W6	Q8T6W6 listeria mo
6	39	79.6	876	16 Q9K854	Q9K854 bacillus ha
7	37	75.5	388	4 Q96LPT	Q96LPT homo sapien
8	36	73.5	114	2 Q8VSV3	Q8VSV3 escherichia
9	36	73.5	114	2 Q8VSV2	Q8VSV2 escherichia
10	36	73.5	114	2 Q8VSV1	Q8VSV1 escherichia
11	36	73.5	114	2 Q8VSV0	Q8VSV0 escherichia
12	36	73.5	281	2 Q9AMB8	Q9AMB8 burkholderi
13	36	73.5	928	16 Q8X8H1	Q8X8H1 escherichia
14	35	71.4	289	16 Q97L97	Q97L97 clostridium
15	35	71.4	471	17 Q972J5	Q972J5 sulfolobus
16	35	71.4	483	17 Q9UZY2	Q9UZY2 pyrococcus

17	35	71.4	980	16 Q9Z7B5	Q9Z7B5 chlamydia p
18	35	71.4	1259	10 Q223Z6	Q223Z6 arabidopsis
19	35	71.4	1259	10 Q231Q2	Q231Q2 arabidopsis
20	35	71.4	1260	10 Q9FHN8	Q9FHN8 arabidopsis
21	35	71.4	1260	10 Q8VZY4	Q8VZY4 arabidopsis
22	35	71.4	1261	10 Q39130	Q39130 arabidopsis
23	34	69.4	319	16 Q67196	Q67196 aquilex ae
24	34	69.4	1285	13 Q9DF50	Q9DF50 xenopus lae
25	34	69.4	1690	5 Q8F9E7	Q8F9E7 drosophila
26	34	69.4	2515	5 Q9W2J8	Q9W2J8 drosophila
27	34	69.4	3616	13 Q3W6V0	Q3W6V0 gallus gall
28	33	67.3	228	5 Q6Z546	Q6Z546 hydra allen
29	33	67.3	243	5 Q8WR84	Q8WR84 tribolium c
30	33	67.3	251	2 Q9ZEX5	Q9ZEX5 buchera ap
31	33	67.3	292	5 Q9HK02	Q9HK02 tribolium c
32	33	67.3	305	16 Q9JZU1	Q9JZU1 neisseria m
33	33	67.3	346	16 Q8VXA3	Q8VXA3 anaerobium sp
34	33	67.3	347	10 Q9S2D0	Q9S2D0 arabidopsis
35	33	67.3	478	3 Q9P8H9	Q9P8H9 ericella
36	33	67.3	606	5 P90573	P90573 plasmidium
37	33	67.3	798	16 Q8U6F8	Q8U6F8 agrobacteri
38	33	67.3	871	2 Q8VPE6	Q8VPE6 thermococci
39	33	67.3	871	16 Q8RHK6	Q8RHK6 thermococci
40	33	67.3	928	16 Q8Z2S2	Q8Z2S2 salmonella
41	33	67.3	1071	16 Q9JXW8	Q9JXW8 neisseria m
42	33	67.3	1079	2 Q959Y9	Q959Y9 neisseria g
43	32.5	65.3	131	16 Q9PHQ4	Q9PHQ4 campylobact
44	32	65.3	120	16 Q8KCM0	Q8KCM0 thermococci
45	32	65.3	198	16 Q8RFB8	Q8RFB8 fusobacteri

ALIGNMENTS

RESULT 1
ID Q45458 PRELIMINARY; PRT; 876 AA.
AC Q45458;
DT 01-NOV-1996 (TREMBL:rel. 01, Created)
DT 01-NOV-1996 (TREMBL:rel. 01, last sequence update)
DT 01-JUN-2002 (TREMBL:rel. 21, last annotation update)
DE DNA polymerase I.
GN POLA.
OC Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RA Altiotta J.M., Pelletier J.L., Ware J.L., Moran L.S., Benner J.S.,
KA Kong H.;
RT "Thermuslabile" hist DNA polymerase I lacks a 3'->5' proofreading
exonuclease activity";
RL Genet. Anal. 12:185-195(1996).
DR EXBL: 033536; MMS2611.1; ..
DR HSSP: P52026; IXL.
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR001098; DNA.pol.
DR InterPro: IPR002298; DNA.pol.
DR InterPro: IPR000513; EXO_N.1.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuclease; 1.
DR PRFam: PF00476; DNA.pol.A; 1.
DR PRFam: PF00068; DNAPOL1.
DR SMART: SM00474; 35EXOC; 1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00482; POLAC; 1.
DR TIGRFAMs: TIGR00593; POLA; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
SQ SEQUENCE 876 AA; 9908 MW; B981BCE95F1651A7 CRC64;

Query Match 98.0%; Score 48; DB 2; Length 876;
 Best Local Similarity 90.9%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLITRKE 11
 ||||| |||
 DB 721 LAONLITRKE 731

RESULT 2
 ID 024675 PRELIMINARY; PRT; 877 AA.
 AC 024675;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase I.
 GN POLG1.
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;
 OC Geobacillus.
 RX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FW10;
 RA Xu P., Leung C.H., Kong R.;
 RT "Isolation and heterologous expression of Bacillus stearothermophilus
 RT DNA polymerase I gene.";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBD databases.
 DR EMBL; U93028; AAB62092.1; -.
 DR HSSP; P52026; 1XWL.
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA pol.
 DR InterPro: IPR002298; DNA pol.
 DR InterPro: IPR00513; Exo_N.I.
 DR InterPro: IPR003584; HNH2.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuclease; 1.
 DR Pfam: PF00476; DNA_POL_A; 1.
 DR PRINTS; PR00468; DNAPOL1.
 DR SMART; SM00474; 3EXOC; 1.
 DR SMART; SM00475; 3EXOC; 1.
 DR SMART; SM00279; HNH2; 1.
 DR SMART; SM00482; POLAC; 1.
 DR TIGRFAMs; TIGR00593; pola; 1.
 DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
 SQ SEQUENCE 877 AA; 99200 MW; 4B9639483FE41DDB CRC64;

Query Match 98.0%; Score 48; DB 2; Length 877;
 Best Local Similarity 90.9%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLITRKE 11
 ||||| |||
 DB 723 LAONLITRKE 733

RESULT 3
 ID 08XIW9 PRELIMINARY; PRT; 866 AA.
 AC 08XIW9;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase I.
 GN POLA OR CPE1994.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 RX NCBI_TaxID=1502;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-13 / TYPE A;
 RX PubMed=1192842;
 RA Shmidt T., Ohtani K., Hirakawa H., Oshima K., Yamashita A.,
 RA Shida T., Ogasawara N., Hattori M., Kunata S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL; AP003192; BAB81700.1; -.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA pol.
 DR InterPro: IPR002298; DNA pol.
 DR InterPro: IPR00513; Exo_N.I.
 DR InterPro: IPR003584; HNH2.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuclease; 1.
 DR Pfam: PF00476; DNA_POL_A; 1.
 DR PRINTS; PR00868; DNAPOL1.
 DR SMART; SM00475; 3EXOC; 1.
 DR SMART; SM00279; HNH2; 1.
 DR SMART; SM00482; POLAC; 1.
 DR TIGRFAMs; TIGR00593; pola; 1.
 DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
 KM Complete proteome.
 SQ SEQUENCE 866 AA; 99118 MW; E7F786F720146615 CRC64;

Query Match 81.6%; Score 40; DB 16; Length 866;
 Best Local Similarity 72.7%; Pred. No. 9.4;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLITRKE 11
 ||||| |||
 DB 711 LSOLNITRKE 721

RESULT 4
 ID 092BF0 PRELIMINARY; PRT; 875 AA.
 AC 092BF0;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE DNA polymerase I.
 GN POLA OR LIN1600.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 RX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Francau L., Buchrieser C., Rusniok C., Amend A.,
 RA Bagueri F., Berche P., Blocker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deboux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
 RA Entian K.-D., Fshil H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurupak H.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schueter T., Simoes N., Tietzer A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL; AL596169; CAC96831.1; -.
 DR ListList; LIN01600; -.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA pol.
 DR InterPro: IPR002298; DNA pol.
 DR InterPro: IPR00513; Exo_N.I.
 DR InterPro: IPR003583; HNH2.
 DR Pfam: PF01367; 5_3_exonuclease; 1.

DR Pfam: PF02739; 5.3-exonuc.N; 1.
 DR Pfam: PF00476; DNA_pol.A; 1.
 DR SMART: SM00278; HnhI; 1.
 DR TIGRFAMs: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 875 AA; 98609 MW; 084F393681BD61D9 CRC64;

Query Match 79.6%; Score 39; DB 16; Length 875;
 Best Local Similarity 72.7%; Pred. No. 15;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLNXRKE 11
 Db 720 LSONLGITRKE 730

RESULT 5
 O8Y6M6 PRELIMINARY; PRT; 875 AA.

AC O8Y6M6;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE DNA polymerase 1.
 GN POLA OR LMO1565.
 OS Listeria monocytogenes.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;

RA SEQUENCE FROM N.A.
 RC STRAIN-EGD-E / SEROVAR 1/2A;
 RX MEDLINE-21537279; PubMed-11679669;
 RA Glaeser P., Frangeul L., Buchrieser C., Rusnock C., Amend A.,
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 Charbit A., Clehouani F., Cove E., de Darvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchud E., Durant L., Dussuguet O.,
 Entian K.-., Eschl H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Iopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurapkak G.,
 Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjati H.,
 RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tietz A.,
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL591979; CAC99643.1; -

DR Listlist; LMO1565; -
 DR InterPro: IPR002562; 3.5-exonuclease.
 DR InterPro: IPR002421; 5.3-exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR000513; Exo_N.I.
 DR InterPro: IPR003584; HNH.2.
 DR Pfam: PF01367; 5.3-exonuclease; 1.
 DR Pfam: PF02739; 5.3-exonuc.N; 1.
 DR Pfam: PF00476; DNA_pol.A; 1.
 DR PRINTS: PR00868; DNAPOLI.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00279; HnhI; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFAMs: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 KW Complete proteome.
 SQ SEQUENCE 875 AA; 98448 MW; 94B541A619D005CF CRC64;

Query Match 79.6%; Score 39; DB 16; Length 875;
 Best Local Similarity 72.7%; Pred. No. 15;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLNXRKE 11

Db 720 LSONLGITRKE 730

RESULT 6
 ID O9K854 PRELIMINARY; PRT; 876 AA.
 AC O9K854;

DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE DNA polymerase I (ec 2.7.7.7).
 GN POLA OR BH153.
 OS Bacillus halodurans.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;

RA SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE-20512582; PubMed-11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001517; BAB06872.1; -
 DR HSSP: P52026; 2BDP.

DR InterPro: IPR002562; 3.5-exonuclease.
 DR InterPro: IPR002421; 5.3-exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR000513; Exo_N.I.
 DR InterPro: IPR003583; HNH.1.
 DR InterPro: IPR003584; HNH.2.
 DR Pfam: PF01367; 5.3-exonuclease; 1.
 DR Pfam: PF02739; 5.3-exonuc.N; 1.
 DR Pfam: PF00476; DNA_pol.A; 1.
 DR PRINTS: PR00868; DNAPOLI.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00279; HnhI; 1.
 DR SMART: SM00278; HnhI; 1.
 DR SMART: SM00279; HnhI; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFAMs: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 KW Transferrase; Nucleotidyltransferase; Complete proteome.
 SQ SEQUENCE 876 AA; 99545 MW; 3CA1D7EBDB689617 CRC64;

Query Match 79.6%; Score 39; DB 16; Length 876;
 Best Local Similarity 72.7%; Pred. No. 15;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLNXRKE 11
 Db 721 LSONLGITRKE 731

RESULT 7
 ID O96LP7 PRELIMINARY; PRT; 388 AA.
 AC O96LP7;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CDNA FLJ25320 f1s, clone TST0267 (Fragment).
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mamalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-TESTIS;
 RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
 RA Hoshida T., Hirooka S., Murakawa K., Takiguchi S., Kusano J.,
 RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
 RA Suzuki Y., Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
 RA Kawakami B., Nagai K., Isogai T., Sugano S.,
 RT "NEO human cDNA sequencing project";
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK058049; BAB71640.1; -
 FT NON_TER 388
 SQ SEQUENCE 388 AA; 43337 MW; C3502218B326F985 CRC64;

Query Match 75.5%; Score 37; DB 4; Length 388;
 Best Local Similarity 63.6%; Pred. NO. 18;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAONLINXKE 11
 ||:|||||
 DB 257 LAQSLNTRQE 267

RESULT 8
 O8VS3 PRELIMINARY; PRT; 114 AA.
 AC O8VS3:
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE DNA polymerase I (Fragment).
 GN POLA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ECOR 2;
 RX MEDLINE=21610577; PubMed=11746762;
 RA Brown E.W., Leclerc J.E., Kotewicz M.L., Cebula T.A.;
 RT "Three R's of bacterial evolution: how replication, repair, and
 recombination frame the origin of species.";
 RL Environ. Mol. Mutagen. 38:248-260(2001).
 DR EMBL; AF359695; AAL59502.1; -
 DR InterPro: IPR001098; DNA_pol.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR SMART; SM00482; POLAC; 1.
 DR PROSITE; PS00447; DNA_POLYMERASE_A; UNKNOWN_1.
 FT NON_TER 114
 FT NON_TER 1
 SQ SEQUENCE 114 AA; 13022 MW; FLFF10DDAF041C85 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 114;
 Best Local Similarity 72.7%; Pred. NO. 8.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLINXKE 11
 ||:|||||
 DB 81 LAQSLNTRQE 91

RESULT 9
 O8VS2 PRELIMINARY; PRT; 114 AA.
 AC O8VS2:
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE DNA polymerase I (Fragment).
 GN POLA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ECOR 46;
 RX MEDLINE=21610577; PubMed=11746762;
 RA Brown E.W., Leclerc J.E., Kotewicz M.L., Cebula T.A.;
 RT "Three R's of bacterial evolution: how replication, repair, and
 recombination frame the origin of species.";
 RL Environ. Mol. Mutagen. 38:248-260(2001).
 DR EMBL; AF359736; AAL59543.1; -
 DR InterPro: IPR001098; DNA_pol.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR SMART; SM00482; POLAC; 1.
 DR PROSITE; PS00447; DNA_POLYMERASE_A; UNKNOWN_1.
 FT NON_TER 114
 FT NON_TER 1
 SQ SEQUENCE 114 AA; 13037 MW; 9DFP10DDAF040155 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 114;
 Best Local Similarity 72.7%; Pred. NO. 8.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLINXKE 11
 ||:|||||
 DB 81 LAQSLNTRQE 91

RESULT 10
 O8VL0 PRELIMINARY; PRT; 114 AA.
 AC O8VL0:
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE DNA polymerase I (Fragment).
 GN POLA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ECOR 44, AND ECOR 47;
 RX MEDLINE=21610577; PubMed=11746762;
 RA Brown E.W., Leclerc J.E., Kotewicz M.L., Cebula T.A.;
 RT "Three R's of bacterial evolution: how replication, repair, and
 recombination frame the origin of species.";
 RL Environ. Mol. Mutagen. 38:248-260(2001).
 DR EMBL; AF359740; AAL59547.1; -
 DR EMBL; AF359741; AAL59548.1; -
 DR InterPro: IPR001098; DNA_pol.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR SMART; SM00482; POLAC; 1.
 DR PROSITE; PS00447; DNA_POLYMERASE_A; UNKNOWN_1.
 FT NON_TER 114
 FT NON_TER 1
 SQ SEQUENCE 114 AA; 13024 MW; 1FP2A6DC2B655B2 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 114;
 Best Local Similarity 72.7%; Pred. NO. 8.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLINXKE 11
 ||:|||||
 DB 81 LAQSLNTRQE 91

RESULT 11
 O8VL8 PRELIMINARY; PRT; 114 AA.
 AC O8VL8:
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE DNA polymerase I (Fragment).
 GN POLA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 NC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA MEDLINE=21610577; PubMed=11746762;
 RX Brown E.W., Leclerc J.E., Kotewicz M.L., Cebula T.A.;
 RT "Three R's of bacterial evolution: how replication, repair, and
 recombination frame the origin of species."
 RL Environ. Mol. Mutagen. 38:248-260(2001).
 DR EMBL: AF359677; AAL59484.1; -
 DR EMBL: AF359678; AAL59485.1; -
 DR EMBL: AF359679; AAL59486.1; -
 DR EMBL: AF359680; AAL59487.1; -
 DR EMBL: AF359681; AAL59488.1; -
 DR EMBL: AF359682; AAL59489.1; -
 DR EMBL: AF359683; AAL59490.1; -
 DR EMBL: AF359684; AAL59491.1; -
 DR EMBL: AF359685; AAL59492.1; -
 DR EMBL: AF359686; AAL59493.1; -
 DR EMBL: AF359687; AAL59494.1; -
 DR EMBL: AF359688; AAL59495.1; -
 DR EMBL: AF359689; AAL59496.1; -
 DR EMBL: AF359690; AAL59497.1; -
 DR EMBL: AF359691; AAL59498.1; -
 DR EMBL: AF359692; AAL59499.1; -
 DR EMBL: AF359693; AAL59500.1; -
 DR EMBL: AF359694; AAL59501.1; -
 DR EMBL: AF359695; AAL59502.1; -
 DR EMBL: AF359696; AAL59503.1; -
 DR EMBL: AF359697; AAL59504.1; -
 DR EMBL: AF359698; AAL59505.1; -
 DR EMBL: AF359699; AAL59506.1; -
 DR EMBL: AF359700; AAL59507.1; -
 DR EMBL: AF359701; AAL59508.1; -
 DR EMBL: AF359702; AAL59509.1; -
 DR EMBL: AF359703; AAL59510.1; -
 DR EMBL: AF359704; AAL59511.1; -
 DR EMBL: AF359705; AAL59512.1; -
 DR EMBL: AF359706; AAL59513.1; -
 DR EMBL: AF359707; AAL59514.1; -
 DR EMBL: AF359708; AAL59515.1; -
 DR EMBL: AF359709; AAL59516.1; -
 DR EMBL: AF359710; AAL59517.1; -
 DR EMBL: AF359711; AAL59518.1; -
 DR EMBL: AF359712; AAL59519.1; -
 DR EMBL: AF359713; AAL59520.1; -
 DR EMBL: AF359714; AAL59521.1; -
 DR EMBL: AF359715; AAL59522.1; -
 DR EMBL: AF359716; AAL59523.1; -
 DR EMBL: AF359717; AAL59524.1; -
 DR EMBL: AF359718; AAL59525.1; -
 DR EMBL: AF359719; AAL59526.1; -
 DR EMBL: AF359720; AAL59527.1; -
 DR EMBL: AF359721; AAL59528.1; -
 DR EMBL: AF359722; AAL59529.1; -
 DR EMBL: AF359723; AAL59530.1; -
 DR EMBL: AF359724; AAL59531.1; -
 DR EMBL: AF359725; AAL59532.1; -
 DR EMBL: AF359726; AAL59533.1; -
 DR EMBL: AF359727; AAL59534.1; -
 DR EMBL: AF359728; AAL59535.1; -
 DR EMBL: AF359729; AAL59536.1; -
 DR EMBL: AF359730; AAL59537.1; -
 DR EMBL: AF359731; AAL59538.1; -
 DR EMBL: AF359732; AAL59539.1; -
 DR EMBL: AF359733; AAL59540.1; -
 DR EMBL: AF359734; AAL59541.1; -
 DR EMBL: AF359735; AAL59542.1; -

DR EMBL: AF359737; AAL59544.1; -
 DR EMBL: AF359738; AAL59545.1; -
 DR EMBL: AF359739; AAL59546.1; -
 DR EMBL: AF359740; AAL59547.1; -
 DR EMBL: AF359741; AAL59548.1; -
 DR EMBL: AF359742; AAL59549.1; -
 DR EMBL: AF359743; AAL59550.1; -
 DR EMBL: AF359744; AAL59551.1; -
 DR EMBL: AF359745; AAL59552.1; -
 DR EMBL: AF359746; AAL59553.1; -
 DR EMBL: AF359747; AAL59554.1; -
 DR EMBL: AF359748; AAL59555.1; -
 DR InterPro: IPR001098; DNA_pol.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR SMART: SM00482; POLA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; UNKNOWN_1.
 FT NON-TER 1
 FT 114 114
 SQ SEQUENCE 114 AA; 13038 MW; 90FF100DAF041C95 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 114;
 Best Local Similarity 72.7%; Pred. No. 8.6;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LAONINIRKE 11
 11:111111
 Db 81 LAROLNIRKE 91

RESULT 12

Q9AMB8 PRELIMINARY; PRT; 281 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)

DE RNA polymerase sigma factor.

GN RPOD.

OS Burkholderia pseudomallei (pseudomonas pseudomallei).

OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;

OC Burkholderia.

NC NCBI_TaxID=28450;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-G9313;

RA Steiner B., Bowen M., Morrill W., Meyer R.;

RT "Random sequencing of Burkholderia pseudomallei strain G9313 for

RT clinical PCR development."

RL Submitted (DMR-2000) to the EMBL/GenBank/DBJ databases.

CC FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES

CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND

CC THEN IS RELEASED (BY SIMILARITY).

CC -1- SIMILARITY: BRINGS TO THE SIGMA-70 FACTOR FAMILY.

DR EMBL: AF326688; AAK0822.1; -

DR HSP: P00579; ISIG.

DR InterPro: IPR000943; Sigma_70.

DR Pfam: PF00140; sigma70; 1.

DR PRINTS: PR00046; SIGMA70FCF.

DR PROSITE: PS00715; SIGMA70_1; 1.

DR PROSITE: PS00716; SIGMA70_2; 1.

KW DNA-binding: DNA-directed RNA polymerase; Sigma factor;

KW Transcription regulation.

SQ SEQUENCE 281 AA; 31352 MW; 4C50367ECB61204 CRC64;

QY 1 LAONINIRKE 11
 11:111111

Db 150 LAOELNIRKE 160

RESULT 13
 Q8X8H1

```

ID 08X8H1 PRELIMINARY; PRT; 928 AA.
AC 08X8H1;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE DNA polymerase I, 3'--> 5' polymerase, 5'--> 3' and 3'--> 5'
DE exonuclease.
GN POLA OR 25398 OR EC54766.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-0157:H7 / EDL933 / ATCC 700927;
RA MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaeser J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-0157:H7 / RIMD 0509952;
RA MEDLINE=21156231; PubMed=11256796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shida T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005617; AAG59052.1; -
DR EMBL; AP002567; BAB38209.1; -
DR InterPro: IPR002562; 3.5_exonuclease.
DR InterPro: IPR002421; 5.3_exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N1.
DR InterPro: IPR003584; HH2_2.
DR Pfam: PF01612; 3.5_exonuclease; 1.
DR Pfam: PF02739; 5.3_exonuc; 1.
DR Pfam: PF00476; DNA_pol_A; 1.
DR PRINTS; PR00868; DNAPOL1.
DR SMART; SM00474; 35EXOC; 1.
DR SMART; SM00475; 53EXOC; 1.
DR SMART; SM00279; HH2; 1.
DR SMART; SM00482; POLA; 1.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Complete proteome.
SQ SEQUENCE 928 AA; 103146 MW; 86431C44859154E CRC64;

Query Match 73.5%; Score 36; DB 16; Length 928;
Best local Similarity 72.7%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLIXRKE 11
ID 097L97 PRELIMINARY; PRT; 289 AA.
AC 097L97;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ABC-type sugar transport system, permease component.

```

```

GN CAC0665.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AVCC 824 / DSM 792 / VKM H-1787;
RA MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hilli J., Wolf Y.I.,
RA Yarusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007582; AAK78642.1; -
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00528; BPD_transp; 1.
DR PROSITE; PS00402; BPD_TRANS_PLN_MEMB; 1.
KW Sugar transport; Complete proteome.
SQ SEQUENCE 289 AA; 32969 MW; 49DB3B014163C2AB CRC64;

Query Match 71.4%; Score 35; DB 16; Length 289;
Best local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 ONLIXRKE 11
ID 0972J5 PRELIMINARY; PRT; 471 AA.
AC 0972J5;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative ABC transporter ATP-binding protein.
GN ST1136.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horiikawa H., Jinno K., Takahashi M.,
RA Sekine M., Baba S.I., Aukai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka K., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kuboh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagih M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermocidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000985; BAB6171.1; -
DR InterPro: IPR003439; ABC_transport.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_TRANSPORT; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW ATP-binding; Hypothetical protein; Complete proteome.
SQ SEQUENCE 471 AA; 53267 MW; 5708B9B1047D562 CRC64;

Query Match 71.4%; Score 35; DB 17; Length 471;
Best local Similarity 70.0%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 AONLIXRKE 11
ID 109 AFNINIPRKE 118

```

Tue Jan 21 14:46:25 2003

us-09-823-649a-7.rspt

Page 7

Search completed: January 15, 2003, 12:36:11
Job time : 33.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 Seconds

(without alignments)
72.584 Million cell updates/sec

Title: US-09-823-649a-7
Perfect score: 49
Sequence: 1 LAQNLNIXRKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	48	98.0	876	1	DPO1_BACST
2	48	98.0	877	1	DPO1_BACCA
3	49	98.0	877	1	DPO1_BACSV
4	38	79.6	880	1	SMX4_SCHMA
5	38	77.6	98	1	SG12_YEAST
6	36	73.5	545	1	P20438 saccharomyc
7	36	73.5	928	1	DPO1_BACLI
8	35	71.4	877	1	DPO1_LACIA
9	35	71.4	877	1	DPO1_LACIC
10	34	69.4	603	1	YA07_SCHPO
11	34	69.4	877	1	DPO1_STRPN
12	34	69.4	2096	1	HP28_DROME
13	33	67.3	2515	1	TUD_DROME
14	33	67.3	161	1	YU57_CLOPE
15	33	67.3	320	1	YC57_ARCRU
16	33	67.3	917	1	YGJ3_YEAST
17	33	67.3	928	1	DPO1_SALTY
18	33	67.3	1071	1	CARB_NEICO
19	32	65.3	1071	1	CARB_NEIMB
20	32	65.3	71	1	RL29_RICCN
21	32	65.3	71	1	RL29_RICCR
22	32	65.3	175	1	Y773_METUA
23	32	65.3	336	1	ASNA_CLOPE
24	31	63.3	497	1	HMS5_DROME
25	31	63.3	173	1	RRY5_CAEEL
26	31	63.3	238	1	Y554_AOUAE
27	31	63.3	342	1	ID12_RICCN
28	31	63.3	342	1	ID12_RICCR
29	31	63.3	415	1	SAHL_SULTO
30	31	63.3	522	1	MSAB_MELIO
31	31	63.3	527	1	GALL_YEAST
32	31	63.3	528	1	TACY_LISTIV
33	31	63.3	780	1	ACON_BOVIN
			780	1	ACON_HUMAN
					P52026 bacillus st
					Q04957 bacillus ca
					Q26603 schistosoma
					P20438 saccharomyc
					P00582 escherichia
					Q04951 lactococcus
					Q02801 lactococcus
					Q10086 schizosacch
					P13252 streptococc
					Q9W75 drosophila
					P25823 drosophila
					Q93md5 clostridium
					Q02901 archaeoglob
					P53148 saccharomyc
					Q94173 salmonella
					Q05959 neisseria g
					Q91xw8 neisseria m
					Q92kx4 rickettsia
					Q92xw3 rickettsia
					Q58183 mechanococ
					Q8x1y5 clostridium
					P18488 drosophila
					Q09355 caenorhabdi
					Q06829 aquifex aeo
					Q92bm7 rickettsia
					Q92x90 rickettsia
					Q97810 sulfolobus
					P14930 n peptide m
					P04885 saccharomyc
					P31831 listeria iv
					P20004 bos taurus
					Q99798 homo sapien

34	31	63.3	850	1	DPO1_ANATH	Q59156 anaerocellu
35	30	61.2	75	1	Y9KD_MAGV	P16607 maguari vir
36	30	61.2	219	1	PIM2_ARCRU	Q27962 archaeoglob
37	30	61.2	238	1	RK5_ODOSI	P49547 odontella s
38	30	61.2	260	1	NU6M_ACACA	Q37371 acanthameob
39	30	61.2	346	1	FTS2_MYCGE	P47539 mycoplasma
40	30	61.2	370	1	MATR_MARPO	P12174 marchantia
41	30	61.2	449	1	HEM1_HELPJ	Q92m5 helicobacte
42	30	61.2	449	1	HEM1_HELPJ	P56125 helicobacte
43	30	61.2	497	1	SR52_HORVU	P49968 hordeum vul
44	30	61.2	497	1	SR52_HORVU	P49969 hordeum vul
45	30	61.2	529	1	TACY_LISTIV	P13128 listeria mo

ALIGNMENTS

```

RESULT 1
DPO1_BACST STANDARD; PRT; 876 AA.
ID DPOL_BACST
AC P52026;
DN 01-OCT-1996 (rel. 34, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL I).
GN POLA OR POL.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RA Riggs M.G., Tudor S., Sivaram M., McDonough S.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Phang S.M., Teo C.Y., Lo E., Wong V.W.;
RL MEDLINE=96001245; PubMed=7557480;
RT "Cloning and complete sequence of the DNA polymerase-encoding gene
RT (batpol1) and characterisation of the Klenow-like fragment from
RT Bacillus stearothermophilus.";
RL Gene 163:65-68(1995).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 297-876.
RA Kiefer J.R., Mao C., Hansen C.J., Bashore C.L., Hogrefe H.H.;
RL MEDLINE=97169447; PubMed=9016716;
RT "Crystal structure of a thermostable Bacillus DNA polymerase I large
RT fragment at 2.1-A resolution.";
RL Structure 5:95-108(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 297-876.
RA Kiefer J.R., Mao C., Berman J.C., Reese L.S.;
RL MEDLINE=98101648; PubMed=9440698;
RT "Visualizing DNA replication in a catalytically active Bacillus DNA
RT polymerase crystal.";
RL Nature 391:304-307(1998).
RN [5]
RP FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
RN EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY. THE OPTIMAL
RN TEMPERATURE FOR ACTIVITY OF 65 DEGREES CELSIUS.
RN [6]
RP CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
RN + (DNA)n.
RN [7]
RP SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
RN [8]
RP SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
RN [9]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN [10]
RP EMBL; LA2111; AAC37139.1; -.

```

FT	VARIANT	695	696	ED -> D (IN STRAIN USED IN REF. 3/4).
FT <td>VARIANT</td> <td>699</td> <td>699</td> <td>A -> P (IN STRAIN USED IN REF. 3/4).</td>	VARIANT	699	699	A -> P (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	699	728	T -> S (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	741	741	Q -> E (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	748	748	A -> R (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	751	751	D -> E (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	790	790	T -> M (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	812	813	SV -> NA (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	816	816	I -> K (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	841	841	P -> S (IN STRAIN USED IN REF. 3/4).
FT	VARIANT	870	870	L -> V (IN STRAIN USED IN REF. 3/4).
FT	CONFLICT	91	91	P -> S (IN REF. 2).
FT	CONFLICT	198	198	E -> K (IN REF. 2).
FT	CONFLICT	352	352	L -> V (IN REF. 2).
FT	CONFLICT	381	382	KW -> NG (IN REF. 2).
FT	CONFLICT	386	386	R -> ACY (IN REF. 2).
FT	CONFLICT	446	446	H -> Q (IN REF. 2).
FT	CONFLICT	448	448	V -> A (IN REF. 1).
FT	CONFLICT	479	480	QP -> HA (IN REF. 2).
FT	CONFLICT	678	678	G -> W (IN REF. 2).
FT	CONFLICT	785	785	S -> T (IN REF. 2).
FT	CONFLICT	814	815	RL -> SV (IN REF. 2).
FT	CONFLICT	828	828	V -> G (IN REF. 2).
FT	CONFLICT	842	842	E -> G (IN REF. 2).
FT	CONFLICT	857	857	T -> A (IN REF. 1).
SEQ	SEQUENCE	876 AA;	98670 MW;	689167A801D543E4 CRC64;
Query Match	Best Local Similarity	98.0%;	Score 48;	DB 1;
Matches	10;	Conservative	0;	Mismatches
Indels	1;	Indels	0;	Gaps
QY	1	LAONINXKKE 11		
Db	721	LAONINXKKE 731		
RESULT 2				
DPOL_BACCA				
ID	DPOL_BACCA	STANDARD;	PRT;	877 AA.
AC	004957;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	DNA polymerase 1 (EC 2.7.7.7) (POL 1).			
GN	POLA.			
OS	Bacillus caldotenax.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1395;			
RN	[1]			
RP	SEQUENCE FROM N.A. AND SEQUENCE OF 1-15.			
RX	M8011NF-935252765; PubMed=8486614;			
RA	Uemori Y., Ishino Y., Fujita K., Asada K., Kato I.;			
RT	"Cloning of the DNA polymerase gene of Bacillus caldotenax and			
RT	characterization of the gene product.";			
RL	J. Biochem. 113:401-410(1993).			
CC	1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE			
CC	EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY.			
CC	1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate			
CC	+ [dNA](N).			
CC	1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.			
CC	1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; D12982; BAA02361.1; -			
DR	PIR; J02056; J02056.			
DR	HSSP; P52026; JXWL.			

DR InterPro: IPR002562; 3-5_exonuclease.
 DR InterPro: IPR002421; 5-3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR00513; Exo_N.I.
 DR InterPro: IPR003584; HNH_2.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5-3_exonuclease; 1.
 DR Pfam: PF02739; 5-3_exonuclease; 1.
 DR PRINTS: PR00868; DNAPOLI.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00279; Hnh2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFS: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 DR Transferrase: DNA-directed DNA polymerase: DNA replication: DNA repair;
 KW Hydrolyase; Exonuclease; DNA-binding.
 FT DOMAIN 1 310 5-3' EXONUCLEASE (BY SIMILARITY).
 FT DOMAIN 311 465 3'-5' EXONUCLEASE (BY SIMILARITY).
 FT DOMAIN 469 877 POLYMERASE (BY SIMILARITY).
 SQ SEQUENCE 877 AA; 99475 MW; DA5FC7F5B6DA6F4 CRC64;

Query Match 98.0%; Score 48; DB 1; Length 877;
 Best local similarity 90.9%; Pred. No. 0.026;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAONLIXRKE 11
 ||||| |||
 DB 722 LAONLIXRKE 732

RESULT 3
 DPOL_BACSU STANDARD; PRT: 880 AA.
 AC 034996;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase I (PC 2.7.7.7) (POL I).
 GN POLA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98048467; PubMed=9387221;
 RA Tlapius A., Galleron N., Sorokin A., Ehrlich S.D.;
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 in the 200 kb rmb-dnaB region."
 RL Microbiology 143:3431-3441(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertolo M.G., Bessieres P., Bolojin A., Borcherdt S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brlonell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Deniot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koeltter P., Konigstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medicine C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Pressac E., Puig P., Punelle B., Rappott G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Toigoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambui R., Wedler E., Wedler H., Weitzenger T.,
 RA Winters P., Wipal A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis."
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY THIS DNA POLYMERASE
 EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
 SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + [DNA](N).
 CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF008220; AAC00350.1; -
 DR EMBL: Z99118; CAH14869.1; -
 DR USSP: P52026; LKWL
 DR Subtilist; Bg12656; polA.
 DR InterPro: IPR002562; 3-5_exonuclease.
 DR InterPro: IPR002421; 5-3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR00513; Exo_N.I.
 DR InterPro: IPR003584; HNH_2.
 DR InterPro: IPR003583; HNH_1.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5-3_exonuclease; 1.
 DR Pfam: PF02739; 5-3_exonuclease; 1.
 DR PRINTS: PR00868; DNAPOLI.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00278; Hnh1; 1.
 DR SMART: SM00279; Hnh2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFS: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 KW Transferrase: DNA-directed DNA polymerase: DNA replication: DNA repair;
 KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
 SQ SEQUENCE 880 AA; 99091 MW; AFR434AFDF26F488 CRC64;

Query Match 79.6%; Score 39; DB 1; Length 880;
 Best local similarity 72.7%; Pred. No. 2.1;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLIXRKE 11
 ||||| |||
 DB 725 LSQNLGTRKE 735

RESULT 4
 SMX4_SCHEMA STANDARD; PRT: 98 AA.
 AC 026603;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Homeobox protein SMX4-4 (Fragment).

GN SMOX-4.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidae;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Puerto Rican;
RX MEDLINE-92399260; PubMed-1356008;
RA Webster P.J., Mansour T.E.;
RT "Conserved classes of homeodomains in Schistosoma mansoni, an early
bilateral metazoan."
RL Mech. Dev. 38:25-32(1992).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M85302; AAA29931.1; -
DR HSPF: P14653; 1872.
DR InterPro: IPR000047; HTH_repressr.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR00031; HTHREPRESSR.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00369; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1
FT DNA_BIND 37 96 HOMEBOX.
SQ SEQUENCE 96 AA; 12219 MW; B8733F4F787F2F04 CRC64;

Query Match 77.6%; Score 38; DB 1; Length 98;
Best Local Similarity 63.6%; Pred. No. 0.3;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Caps 0;

QY 1 LAONLINIRKE 11
| | | | | :
| | | | | :
Db 70 LAONLINIRKE 80

RESULT 5
CG12_YEAST STANDARD; PRT; 545 AA.
AC P20438;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE G1/S-specific cyclin CLN2.
GN CLN2 OR YPL256C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89345642; PubMed-2569741;
RA Hadwiger J.A., Wittenberg C., Richardson H.E., de Barros Lopes M.,
RT Reed S.I.;
RL "A family of cyclin homologs that control the G1 phase in yeast."
RL Proc. Natl. Acad. Sci. U.S.A. 86:6255-6259(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-90326560; PubMed-2197605;
RA Hadwiger J.A., Reed S.I.;
RT "Nucleotide sequence of the Saccharomyces cerevisiae CLN1 and CLN2

RT genes."
RL Nucleic Acids Res. 18:4025-4025(1990).
RN [3]
RP REVISIONS.
RA Wittenberg C., Chapman-Shimshoni D.;
RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288c / AB972;
RX MEDLINE-97313271; PubMed-9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Bortstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., Dipolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floehel M., Fortin N., Friesen J.D., Fritz C., Goleau A.,
RA Hall J., Hedling U., Heumann K., Hilbert H., Hillier L.,
RA Hunkeler-Smith S., Hyman R., Johnston M., Kalan S., Kleene K.,
RA Komp C., Kurt O., Lashkari D., Jew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Moses H.-W., Mitalpali S., Mostl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petal F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherens B., Schramm S., Schroeder M., Sedzu A.M., Tellelin H.,
RA Urrestarazu U.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambolt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zolner A., Vo D.H., Hant J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."
RL Nature 387:103-105(1997).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION. INTERACTS WITH THE CDC28 PROTEIN KINASE TO
CC FORM MPF.
CC -1- DEVELOPMENTAL STAGE: CLN1 AND CLN2 MRNAS FLUCTUATE PERIODICALLY IN
CC THE CELL CYCLE, PEAKING IN G1 PHASE.
CC -1- MISCELLANEOUS: A DOMINANT MUTATION IN CLN2 GENE (CLN2-1), ADVANCES
CC THE G1- TO S-PHASE TRANSITION IN CYCLING CELLS AND IMPAIRS THE
CC ABILITY OF CELLS TO ARREST IN G1 PHASE IN RESPONSE TO EXTERNAL
CC SIGNALS.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. STRONGEST TO OTHER
CC G1/S CYCLINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M33265; AAA65725.1; -
DR EMBL: Z73612; CAAG7982.1; -
DR PIR: B33289; CYBNC2.
DR SGD: S0006177; CLN2.
DR InterPro: IPR004366; Cyclin.
DR Pfam: PF00134; CYCLIN; 1.
DR SMART: SM00385; CYCLIN; 1.
DR PROSITE: PS00292; CYCLINS; 1.
KW Cyclin; Cell cycle; Cell division.
SQ SEQUENCE 545 AA; 61696 MW; D6426B94E040E960 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 545;
Best Local Similarity 70.0%; Pred. No. 5.2;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Caps 0;

QY 1 LAONLINIRK 10
| | | | | :
| | | | | :
Db 367 IASNLINIRK 376

RESULT 6
DPOI_ECOLI STANDARD; PRT; 928 AA.
AC P00582;
DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE POLA polymerase I (PC 2.7.7.7) (POL 1).
CN POLA OR RPSA OR B3863.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-82120160; PubMed-6276402;
RA Joyce C.M., Kelley W.S., Grindley N.D.F.;
RT "Nucleotide sequence of the Escherichia coli polA gene and primary
RT structure of DNA polymerase I.";
RL J. Biol. Chem. 257:1958-1964(1982).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-93347969; PubMed-8346018;
RA Plunkett G., III, Burland V.D., Daniels D.L., Hattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes.";
RL Nucleic Acids Res. 21:3391-3396(1993).
RN [13]
RP SEQUENCE OF 918-928 FROM N.A.
RC STRAIN-K12;
RX MEDLINE-83056713; PubMed-6183253;
RA Joyce C.M., Grindley N.D.;
RT "Identification of two genes immediately downstream from the polA
RT gene of Escherichia coli.";
RL J. Bacteriol. 152:1211-1219(1982).
RN [14]
RP SEQUENCE OF 284-350 FROM N.A.
RX MEDLINE-83189123; PubMed-6302278;
RA Kelley W.S., Joyce C.M.;
RT "Genetic characterization of early amber mutations in the Escherichia
RT coli polA gene and purification of the amber peptides.";
RL J. Mol. Biol. 164:529-560(1983).
RN [15]
RP COMPOSITION, AND PARTIAL SEQUENCE.
RX MEDLINE-82120161; PubMed-7035456;
RA Brown W.E., Slump K.H., Kelley W.S.;
RT "Escherichia coli DNA polymerase I. Sequence characterization and
RT secondary structure prediction.";
RL J. Biol. Chem. 257:1965-1972(1982).
RN [16]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-85137890; PubMed-3863192;
RA Ollis D.L., Brick P., Hamlin R., Xuong N.G., Steitz T.A.;
RT "Structure of large fragment of Escherichia coli DNA polymerase I
RT complexed with dTMP.";
RL Nature 313:762-766(1985).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-91114709; PubMed-1989886;
RA Heese L.S., Steitz T.A.;
RT "Structural basis for the 3'-5' exonuclease activity of Escherichia
RT coli DNA polymerase I: a two metal ion mechanism.";
RL EMBO J. 10:25-33(1991).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-93227044; PubMed-846987;
RA Heese L.S., Derbyshire V., Steitz T.A.;
RT "Structure of DNA polymerase I Klenow fragment bound to duplex DNA.";
RL Science 260:352-355(1993).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-94083412; PubMed-8260491;
RA Heese L.S., Friedman J.M., Steitz T.A.;
RT "Crystal structures of the Klenow fragment of DNA polymerase I
RT complexed with deoxynucleoside triphosphate and pyrophosphate.";
RL Biochemistry 32:14095-14101(1993).

RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-98181033; PubMed-9514742;
RA Brattigam C.A., Steitz T.A.;
RT "Structural principles for the inhibition of the 3'-5' exonuclease
RT activity of Escherichia coli DNA polymerase I by phosphorothioates.";
RL J. Mol. Biol. 277:363-377(1998).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-99109320; PubMed-9888810;
RA Brattigam C.A., Sun S., Piccirilli J.A., Steitz T.A.;
RT "Structures of normal single-stranded DNA and deoxyribo-3'-5'-
RT phosphorothioates bound to the 3'-5' exonuclease active site of
RT DNA polymerase I from Escherichia coli.";
RL Biochemistry 38:646-704(1999).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.08 ANGSTROMS) OF KLENOW FRAGMENT.
RX MEDLINE-20056229; PubMed-10588690;
RA Teplova M., Wallace S.T., Tereshko V., Minusov G., Symons A.M.,
RA Cook P.D., Manoharan M., Eji M.;
RT "Structural origins of the exonuclease resistance of a zwitterionic
RT RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:14240-14245(1999).
RN [13]
RP STRUCTURE BY NMR OF 728-777.
RX MEDLINE-93183012; PubMed-8442659;
RA Mullen G.P., Vaughn J.H., Jr., Milvan A.S.;
RT "Sequential proton NMR resonance assignments, circular dichroism, and
RT structural properties of a 50-residue substrate-binding peptide from
RT DNA polymerase I.";
RL Arch. Biochem. Biophys. 301:174-183(1993).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY. IT IS ABLE
CC TO UTILIZE NICKED CIRCULAR DUPLEX DNA AS A TEMPLATE AND CAN
CC UNWIND THE PARTIAL DNA STRAND FROM ITS TEMPLATE.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA)(N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -1- DATABASE: NAME:worthington enzyme manual;
CC WWW="http://www.worthington-biochem.com/manual/DNAP.html".
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).
DR EMBL: V00417; CAA24607.1;
DR EMBL: L19201; AAO02498.1;
DR EMBL: AAC000461; AAC76881.1;
DR EMBL: J01663; AAA24402.1;
DR EMBL: J01664; AAA24404.1;
DR PIR: A00718; DJET1.
DR PIR: S40809; S40809.
DR DDB: 10P1; 16-OCT-87.
DR DDB: 1KHD; 22-JUN-94.
DR DDB: 1KFS; 25-FEB-98.
DR DDB: 1KLN; 30-NOV-94.
DR DDB: 1KRP; 25-FEB-98.
DR DDB: 1KSP; 25-FEB-98.
DR DDB: 2KPN; 11-NOV-98.
DR DDB: 2KF7; 11-NOV-98.
DR DDB: 2KEM; 16-FEB-99.
DR DDB: 2KE2; 16-FEB-99.
DR DDB: 1QSL; 30-JUN-99.
DR DDB: 1DSY; 02-DIC-99.
DR DDB: 1DSD; 02-DIC-99.
DR DDB: 1DHF; 02-DIC-99.
DR SWISS-PROT: P00582; COIL.
DR ECDDBASE: F113.0; 6TH EDITION.

DR Ecogene; EG10746; polA.
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR00513; Exo_N_1.
 DR InterPro: IPR003583; HHH_1.
 DR InterPro: IPR003584; HHH_2.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF01612; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuclease; 1.
 DR PRINTS: PR00868; DNAPOL1.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00278; 53EXOC; 1.
 DR SMART: SM00279; HHH2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFS: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 DR Transferrase: DNA-directed DNA polymerase; DNA replication; DNA repair;
 KM Hydrolase; Nuclease; Exonuclease; DNA-binding; 3D-structure;
 KW Complete proteome.
 SQ SEQUENCE 928 AA; 10318 MW; DAELC448A59030C CRC64;

Query Match 73.5%; Score 36; DB 1; Length 928;
 Best Local Similarity 72.7%; Pred. No. 9.4;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLIXKE 11
 ||:|||||
 Db 773 LAROLNPKRE 783

RESULT 7
 DPOL_LACIA STANDARD; PRT; 877 AA.
 AC 09CDS1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase I (EC 2.7.7.7) (POL I).
 GN POLA OR IL2142.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Talion O., Malarme K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.
 RL Genome Res. 11:731-753(2001).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA](N).
 CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: AE006442; AAK06240.1; -
 DR HSSP: P52026; 1XWL.
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR00513; Exo_N_1.
 DR InterPro: IPR003584; HHH_2.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuclease; 1.
 DR PRINTS: PR00868; DNAPOL1.
 DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00278; 53EXOC; 1.
 DR SMART: SM00279; HHH2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFS: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 DR Transferrase: DNA-directed DNA polymerase; DNA replication; DNA repair;
 KM Hydrolase; Nuclease; Exonuclease; DNA-binding; Complete proteome.
 SQ SEQUENCE 877 AA; 98732 MW; A5C2B592F98FB2 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 877;
 Best Local Similarity 63.6%; Pred. No. 14;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLIXKE 11
 ||:|||||
 Db 722 LARNIGIRKD 732

RESULT 8
 DPOL_LACIA STANDARD; PRT; 877 AA.
 AC 032801;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase I (EC 2.7.7.7) (POL I).
 GN POLA.
 OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MG1363;
 RX MEDLINE=97369814; PubMed=9246255;
 RA Duwat P., Cochut A., Ehrlich S.D., Gross A.;
 RT "Characterization of Lactococcus lactis UV-sensitive mutants obtained
 RT by ISSI transposition."
 RL J. Bacteriol. 179:4473-4479(1997).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA](N).
 CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: U78771; AAB64184.1; -
 DR HSSP: P52026; 1XWL.
 DR InterPro: IPR002562; 3_5_exonuclease.
 DR InterPro: IPR002421; 5_3_exonuclease.

DR InterPro; IPR001098; DNA_pol.
DR InterPro; IPR002298; DNA_pol.
DR InterPro; IPR00513; Exo_N.I.
DR InterPro; IPR003584; HNH_2.
DR Pfam; PF00476; DNA_pol_A.1.
DR Pfam; PF01367; 5_3_exonuclease_1.
DR PRINTS; PR00868; DNAPOL1.
DR SMART; SM00474; 35EXOC; 1.
DR SMART; SM00475; 53EXOC; 1.
DR SMART; SM00279; HNH2; 1.
DR SMART; SM00482; POLAC; 1.
DR TIGRfams; TIGR00593; polA; 1.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Transferrase: DNA-directed DNA polymerase; DNA replication; DNA repair;
KW Hydrolase: exonuclease; DNA-binding.
SQ SEQUENCE 877 AA: 98895 MW: 580355589620 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 877;
Best local Similarity 63.6%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLIXRKE 11
DB 722 LARNLGTTRKD 732

RESULT 9
ID YAO7 SCHEP STANDARD; PRT: 603 AA.
AC Q10086;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative transcriptional regulatory protein C11D3.07c.
GN SPAC1103.07c
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouris J., Peat N., Hayes J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth W., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Nibbel D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymopre B.,
RA Wellens J., Vansteelt E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrs H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wandut R., Funnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaune V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINDING

CLUSTER DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC
DR EMBL; Z68166; CA92308.2; -
DR HSSP; P08657; TOLD.
DR InterPro; IPR001138; Fungi_TRN.
DR Pfam; PF00172; zn_c1us; 1.
DR SMART; SM00663; GAL4; 1.
DR PROSITE; PS00463; ZN2_C16_FUNGAL_2; 1.
DR PROSITE; PS50048; ZN2_C16_FUNGAL_2; 1.
KW Hypothetical protein; transcription regulation; DNA-binding;
KW Nuclear protein; zinc; Metal-binding.
FT DNA-BIND 7 34
SQ SEQUENCE 603 AA: 69815 MW: 5408089957FF415 CRC64;

Query Match 69.4%; Score 34; DB 1; Length 603;
Best local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LAONLIXRKE 11
DB 282 MAQNLNLRKQ 292

RESULT 10
ID DPOL STRPN STANDARD; PRT: 877 AA.
AC P13252;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL I).
GN POLA OR SP0032.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89139504; PubMed=2537309;
RA Lopez P., Martinez S., Diaz A., Espinosa M., Lacks S.A.;
RT "Characterization of the polA gene of Streptococcus pneumoniae and
RT from Escherichia coli and phage T7.";
RL J. Biol. Chem. 264:4255-4263(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umrigar L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Ufferback T.R., Hansen G.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA] (N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04479; AAA26954.1; -
DR EMBL: AE007320; AAK74222.1; -
DR PIR: A32949; A32949.
DR HSSP: P00582; 1KFS.
DR TIGR: SP0032; -
DR InterPro: IPR002562; 3_5-exonuclease.
DR InterPro: IPR002421; 5_3-exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N.I.
DR InterPro: IPR003584; HNH.2.
DR Pfam: PF00476; DNA_pol_A.1.
DR Pfam: PF01367; 5_3-exonuclease; 1.
DR Pfam: PF02739; 5_3-exonuc. N; 1.
DR PRINTS: PR00868; DNAPOLI.
DR SMART: SM00474; 53EXOC; 1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; HH2; 1.
DR SMART: SM00482; POLAC; 1.
DR TIGRFAWS: TIGR00593; polA; 1.
DR PROSITE: PS00447; DNA-POLYMERASE_A; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
FT CONFLICT 32 32..* V -> A (IN REF. 1).
FT CONFLICT 109 109 D -> E (IN REF. 1).
FT CONFLICT 292 292 V -> M (IN REF. 1).
FT CONFLICT 298 300 SES -> AEG (IN REF. 1).
FT CONFLICT 338 338 V -> A (IN REF. 1).
FT CONFLICT 344 344 K -> Q (IN REF. 1).
FT CONFLICT 493 493 M -> V (IN REF. 1).
FT CONFLICT 697 697 D -> N (IN REF. 1).
FT CONFLICT 790 790 R -> A (IN REF. 1).
SQ SEQUENCE 877 AA; 99252 MW; 262186594586913 CRC64;
Query Match 69.4%; Score 34; DB 1; Length 877;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 LAONINIXRK 11
DB 722 LSNIGISRKE 732
RESULT 11
BP28_DROME STANDARD; PRT; 2096 AA.
AC Q9VM75;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein CG10805.
GN CG10805.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Burton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Sinton R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Aghayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,
RA Ballew R.M., Basu A.K., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Henos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borikova D., Borchen M.R., Bouck J., Brokstein P., Bottler P.,
RA Burris K.C., Busan P.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dong L.E., Downes M., Duquenois S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Mitsuhashi N.Y., Mobarry C., Morris J., Moshell A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svrtk R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weisslock G.M., Weisslock J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
CC -1- SIMILARITY: BM1005 TO THE BAP28 FAMILY.
CC -1- SIMILARITY: BM1005 TO THE BAP28 FAMILY.
CC -1- SIMILARITY: BM1005 TO THE BAP28 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003615; AAF52447.2; -
DR FlyBase: FBgn0031864; CG10805.
DR InterPro: IPR000357; HNAI_FRepeat.
DR PROSITE: PS50077; HNAI_REPEAT; FAMILR_NHC.
KW Hypothetical protein.
FT REPEAT 2058 2094 HNAI
SQ SEQUENCE 2096 AA; 237217 MW; 367B36C7CE6DF62C CRC64;
Query Match 69.4%; Score 34; DB 1; Length 2096;
Best Local Similarity 60.0%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 LAONINIXRK 10
DB 510 LONINIXRK 519
RESULT 12
TUD_DROME STANDARD; PRT; 2515 AA.
AC P25823;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Maternal tudor protein.
GN TUD.
OS Drosophila melanogaster (Fruit fly).

```



```
CC RA STRAIN-13 / Type A:
CC RX PubMed-11792842:
CC RA Shimizu T., Ohnani K., Hirakawa H., Ohshtina K., Yamashita A.,
CC RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.:
CC RT "Complete genome sequence of Clostridium perfringens, an anaerobic
CC RT flesh-eater.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -1-SIMILARITY: no C.perfringens pfr404 OMP6.
CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC or send an email to license@isdb-sib.ch).
CC
DR EMBL; AP003515; BAB62450.1; -
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 161 AA; 19180 MW; DBE5BBA1BB4CBDE CRC64;
CC
Query Match
Best Local Similarity 67.3%; Score 33; DB 1; Length 161;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 LAQNLTNRKE 11
DQ |||::|||
Db 6 LAELNVTFRSE 16
CC
RESULT 14
YC57_ARCFU STANDARD: PRT; 320 AA.
AC 029011:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DS Hypothetical protein AF1257.
GS AF1257.
OS Archaeoglobus fulgidus.
NC Archaea: Euryarchaeota: Archaeoglobi: Archaeoglobales:
OC Archaeoglobacae: Archaeoglobus.
OX NCBI_TaxID=2234;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE-98049343; PubMed-9389475;
RA Kleck H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Ornaghi M., Hickey E.K., Peterson J.D.,
RA Richardson D.H., Kertevadje A.R., Graham D.E., Kyriades N.C.,
RA Reischman R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Klinkens E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glöck A., Zhou L.,
RA Overbeek R., Goehayre J.D., Weidman J.F., McDonald I., Utterback T.,
RA Cotton M.P., Spriggs T., Artlich P., Keane B.P., Sykes S.M.,
RA Sadov P.W., d'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.T., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -1-SIMILARITY: BELONGS TO THE UPF0020 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC or send an email to license@isdb-sib.ch).
CC
EMBL; AE001017; AAB89985.1; -
XR TIGR; AF1257.-.
```

/note="F421Y in AAY91930"

W0200017330-A1.

30-MAR-2000.

14-JAN-1999; 99WO-1B00146.

21-SEP-1998; 98US-0157397.

(SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.

Hong G, Huang W;

WPI; 2000-283563/24.

N-PSDB; AAA08545.

PT New DNA polymerase, useful particularly in DNA sequencing, incorporates all four dideoxynucleotide triphosphate terminators with about equal efficiency

Disclousure; Page 82-90; 109pp; English.

CC The present sequence shows a wild type *Bacillus stearothermophilus* DNA polymerase that is modified (see AAY91930) so that during DNA sequencing, it effectively incorporates all four fluorescent dye-labeled dideoxynucleotide triphosphates (ddNTP), or their analogs, and reduces the discrimination, observed with the unmodified enzyme, for incorporation of labeled ddCTP and ddATP, is new. The modified DNA polymerase, has proofreading 3'-5' exonuclease activity during DNA sequencing from a template and the polymerase excises mismatched nucleotides from the 3' terminus at a faster rate than it removes correctly matched nucleotides. The polymerase is used particularly for CC DNA sequencing, also, for filling in 5'-overhangs, synthesis of DNA CC probes by random primer labeling and in site-directed mutagenesis, CC (claimed).

SQ Sequence 587 AA;

Query Match 98.0%; Score 48; DB 21; Length 587;

Best Local Similarity 90.9%; Pred. No. 0.2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNIXRKE 11

Db 432 LAONLNITRKE 442

RESULT 7

ID AAE11060 standard; Protein: 587 AA.

XX AAE11060;

DT 18-DEC-2001 (first entry)

DE *Bacillus stearothermophilus* (Bst) DNA polymerase 1 protein fragment.KW Reverse transcription; DNA polymerase 1; *Bacillus stearothermophilus*;

KM Bst; enzyme; amplification; hybridisation.

OS *Bacillus stearothermophilus*.

PN W0200164954-A1.

PD 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US06571.

PR 02-MAR-2000; 2000US-0517871.

PA (PROM-) PROMEGA CORP.

PI Gu T, Huang F, Hartnett JR;

DR WPI; 2001-589872/66.

DR N-PSDB; AAD18459.

PT Reverse transcription of RNA using Tvu and Bst DNA polymerase, comprises combining the polymerase, RNA containing sample, a primer and reaction buffer having magnesium ions and reacting the mixture to obtain cDNA

Claim 29; Fig 10; 118pp; English.

CC The invention relates to reverse transcription of RNA templates using CC thermostable DNA polymerase such as *Thermococcus* *vulgaris* (Tvu) and CC *Bacillus stearothermophilus* (Bst) DNA polymerase. Tvu and Bst DNA CC polymerase are useful for reverse transcription reactions which allow CC repetitive heating/cooling cycles without the requirement of fresh enzyme at each cooling step. The invention contemplates single-reaction RT-PCR CC wherein reverse transcription and amplification are performed in a single CC continuous procedure. Primers, template, nucleoside triphosphates, CC appropriate reaction buffer having magnesium ions, reaction conditions, CC and polymerase are used in the PCR process, which involves denaturation CC of target DNA, hybridisation of primers and synthesis of complementary CC strands. The present sequence is *Bacillus stearothermophilus* (Bst) DNA CC polymerase 1 protein fragment which lacks 5' to 3' exonuclease activity.

SQ Sequence 587 AA;

Query Match 98.0%; Score 48; DB 22; Length 587;

Best Local Similarity 90.9%; Pred. No. 0.2;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNIXRKE 11

Db 432 LAONLNITRKE 442

RESULT 8

ID AAR80140 standard; Protein: 588 AA.

XX AAR80140;

DT 30-MAY-1996 (first entry)

DE *B. stearothermophilus* DNA polymerase encoded by clone bst3.

KW *Bacillus stearothermophilus*; thermostable DNA polymerase; PCR; primer; amplification; probe; *Bacillus caldwellii*; exonuclease; recombination; sequencing; polymerase chain reaction.

OS Synthetic.

PN W09527067-A1.

PD 12-OCT-1995.

PF 30-MAR-1995; 95WO-US04080.

PR 24-FEB-1995; 95US-0394232.

PR 01-APR-1994; 94US-022612.

PR 16-SEP-1994; 94US-0307410.

PA (GENP-) GEN-PROBE INC.

PI Riggs MG, Sivaram M, Tudor SD;

DR WPI; 1995-373510/48.

DR N-PSDB; AAT04803.

PT *Bacillus stearothermophilus* DNA encoding thermostable polymerase for nucleic acid sequencing and amplification

PS Claim 9; Page 87-90; 135pp; English.
 XX
 CC This is the amino acid sequence of the modified *Bacillus*
 CC *stearotherophilus* thermostable DNA polymerase encoded by the clone Bst3
 CC which contains a deletion in the 5'-3' exonuclease domain. The corresp.
 CC gene was constructed by digestion and recombination of the full length
 CC gene (AA04801) to remove a 195 bp *HincII*-*PvuII* fragment from the region
 CC encoding the 5'-3' exonuclease. The deleted fragment removes 65 amino
 CC acids (residues 178-242 of the protein) including 2 Gly residues though
 CC to correspond to 2 amino acids in the *E. coli* DNA polymerase I necessary
 CC for 5'-3' exonuclease activity. The modified DNA polymerase can be used
 CC in primer extension reactions such as DNA sequencing or polymerase chain
 CC reaction.
 XX
 SQ Sequence 588 AA;
 Query Match 98.0%; Score 48; DB 16; Length 588;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAQNLTNRKE 11
 DB 433 LAQNLTNRKE 443
 RESULT 9
 ID AAY91930
 XX AAY91930 standard; Protein; 588 AA.
 AC AAY91930;
 XX
 DT 19-JUL-2000 (first entry)
 XX
 DE B. *stearotherophilus* 320 modified DNA polymerase.
 XX
 KW Modified DNA polymerase; DNA sequencing; 3' to 5' exonuclease activity;
 KW random primer labeling; site-directed mutagenesis.
 XX
 OS Synthetic.
 OS *Bacillus stearotherophilus*.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 342..344
 FT /note= "Wild type sequence is Leu-Glu-Glu"
 FT Misc-difference 422
 FT /note= "T422X substitution from wild type"
 XX
 PN WO200017330-A1.
 XX
 PD 30-MAR-2000.
 XX
 PF 14-JAN-1999; 99WO-IB00146.
 XX
 PR 21-SEP-1998; 98US-0157397.
 XX
 PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.
 XX
 PI Hong G, Huang W;
 XX
 DR WI; 2000-283563/24.
 DR N-PSDB; AAA08544.
 XX
 PT New DNA polymerase, useful particularly in DNA sequencing, incorporates
 PT all four dideoxynucleotide triphosphate terminators with about equal
 PT efficiency
 XX
 PS Claim 7; Page 95-103; 109pp; English.
 XX
 CC The present sequence shows a modified *Bacillus stearotherophilus*
 CC DNA polymerase that, during DNA sequencing, effectively
 CC incorporates all four fluorescent dye-labeled dideoxynucleotide
 CC triphosphates (ddNTP), or their analogs, and reduces the discrimination,
 CC observed with the unmodified enzyme, for incorporation of labeled

CC ddNTP and ddATP, is new. The modified DNA polymerase, has proofreading
 CC 3'-5' exonuclease activity during DNA sequencing from a template and the
 CC polymerase excises mismatched nucleotides from the 3' terminus at a
 CC faster rate than it removes correctly matched nucleotides. The
 CC polymerase is used particularly for DNA sequencing, also for filling in
 CC 5'-overhangs, synthesis of DNA probes by random primer labeling and in
 CC site-directed mutagenesis, (claimed).
 XX
 SQ Sequence 588 AA;
 Query Match 98.0%; Score 48; DB 21; Length 588;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAQNLTNRKE 11
 DB 433 LAQNLTNRKE 443
 RESULT 10
 ID AAE11058
 XX AAE11058 standard; Protein; 588 AA.
 AC AAE11058;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Thermoactinomyces vulgaris Tvu DNA polymerase T289M truncated mutant.
 XX
 KW Reverse transcription; DNA polymerase; Thermoactinomyces vulgaris; Tvu;
 KW enzyme; amplification; hybridisation; mutant; mutagen.
 XX
 OS Synthetic.
 OS Thermoactinomyces vulgaris.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 1
 FT /note= "Encoded by AVG"
 XX
 PN WO200164954-A1.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06571.
 XX
 PR 02-MAR-2000; 2000US-0517871.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Gu T, Huang F, Hartnett JK;
 XX
 DR WI; 2001-589872/66.
 DR N-PSDB; AAD18445.
 XX
 PT Reverse transcription of RNA using Tvu and Bst DNA polymerase,
 PT comprises combining the polymerase, RNA containing sample, a primer and
 PT reaction buffer having magnesium ions and reacting the mixture to
 PT obtain cDNA
 XX
 PS Claim 7; Fig 6; 118pp; English.
 XX
 CC The invention relates to reverse transcription of RNA templates using
 CC thermostable DNA polymerase such as Thermoactinomyces vulgaris (Tvu) and
 CC *Bacillus stearotherophilus* (Bst) DNA polymerase. Tvu and Bst DNA
 CC polymerase are useful for reverse transcription reactions which allows
 CC repetitive heating/cooling cycles without the requirement of fresh enzyme
 CC at each cooling step. The invention contemplates single-reaction RT-PCR
 CC wherein reverse transcription and amplification are performed in a single
 CC continuous procedure. Primers, template, nucleoside triphosphates,
 CC appropriate reaction buffer having magnesium ions, reaction conditions,
 CC and polymerase are used in the PCR process, which involves denaturation
 CC of target DNA, hybridisation of primers and synthesis of complementary
 CC strands. The present sequence is Thermoactinomyces vulgaris Tvu DNA

CC polymerase T289M truncated mutant. This Tvu T289M mutant is obtained
CC by deleting 288 amino acids from the N-terminal end of wild-type T.
CC vulgaris Tvu DNA polymerase.
XX

SO Sequence 588 AA;

Query Match 98.0%; Score 48; DB 22; Length 588;
Best Local Similarity 90.9%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAQNINIRKE 11
|||||
DB 433 LAQNINIRKE 443

RESULT 11
AAB85951
ID AAB85951 standard; Protein: 588 AA.

XX AAB85951;

DT 30-NOV-2001 (first entry)

DE Tvu DNA polymerase deletion mutant T289M.

XX Thermotable; DNA polymerase; hyperthermophilic; DNA synthesis;
KW Tvu DNA polymerase; nucleic acid amplification; reverse transcription;
KM deletion mutant; T289M.

OS Thermotomomyces vulgaris.

XX WO200164838-A2.

PN 07-SEP-2001.

PD 28-FEB-2001; 2001WO-US06436.

PR 02-MAR-2000; 2000US-0511439.

PA (PROM-) PROMEGA CORP.

PI Gu T, Huang F, Hartnett JR;

DR WPI: 2001-557768/62.

DR N-PSDB: AAH47392.

PT Novel thermostable DNA polymerase derived from Thermotomomyces
PT vulgaris for use in many recombinant DNA techniques, including
PT amplification, reverse transcription and sequencing reactions
XX
PS Claim 13; Fig 8; 113pp; English.

CC The invention provides compositions comprising thermostable DNA
CC polymerases derived from hyperthermophilic eubacteria, in particular,
CC a purified and isolated Thermotomomyces vulgaris (Tvu) DNA polymerase
CC having DNA synthesis activity. It is useful for determining the
CC nucleotide base sequence of a DNA molecule. Tvu DNA polymerase can be
CC used for processes of high temperature nucleic acid amplification and
CC sequencing without substantial loss of DNA synthesis activity. The enzyme
CC can be used to perform high temperature reverse transcription in the
CC absence of manganese ions and in high throughput robotically-manipulated
CC procedures because greater enzymatic stability is retained at room
CC temperature. Using the Tvu DNA polymerase, reverse transcription
CC reactions can be conducted at increased temperatures so that secondary
CC structure is removed or limited. The present sequence represents the
CC amino acid sequence of a 5'-3' exonuclease deletion mutant form of
CC Tvu DNA polymerase called T289M.

XX Sequence 588 AA;

Query Match 98.0%; Score 48; DB 22; Length 588;
Best Local Similarity 90.9%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAQNINIRKE 11
|||||
DB 433 LAQNINIRKE 443

RESULT 12
AAW30094
ID AAW30094 standard; Protein: 589 AA.

XX AAW30094;

DT 14-APR-1998 (first entry)

DE Bacillus stearothermophilus DNA polymerase (peptides 1-589).

XX Bacillus stearothermophilus; DNA polymerase; DNA amplification;
KW DNA replication; exonuclease; proof reading.

OS Bacillus stearothermophilus.

PN JP09220087-A.

PD 26-AUG-1997.

PE 17-MAY-1996; 96JP-0160402.

PR 03-MAY-1996; 96US-0642684.

PR 18-OCT-1995; 95US-0544643.

PA (HONG/) HONG G F.

PA (HUAN/) HUANG W.

PA (ZHAI/) ZHAI F.

PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.

PI Hong GF, Huang W, Zhai F;

DR WPI: 1997-474304/44.

DR N-PSDB: AAT90820.

PS Claim 9; Pages 29-30; 32pp; Japanese.

CC This is the protein sequence (589 amino acids) of the Bacillus
CC stearothermophilus strain DNA polymerase. The heat stable DNA polymerase
CC can be used in a method for the replication of a DNA chain. It can also
CC be used in a method for the determination of the sequence of DNA chain.
CC As the DNA polymerase has a proof reading 3'-5' exonuclease activity,
CC possible mismatches of base pairing that occurs in current methods of DNA
CC sequence determination of a DNA clone can be avoided.

XX Sequence 589 AA;

Query Match 98.0%; Score 48; DB 18; Length 589;
Best Local Similarity 90.9%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAQNINIRKE 11
|||||
DB 434 LAQNINIRKE 444

RESULT 13
AAR80141
ID AAR80141 standard; Protein: 590 AA.

XX AAR80141;

DT 30-MAY-1996 (first entry)

DE B.steartothermophilus DNA polymerase encoded by clone Bst4.
 XX
 KM Bacillus steartothermophilus; thermostable DNA polymerase; PCR; primer;
 KM amplification; probe; Bacillus caldotenax; exonuclease; recombination;
 KW sequencing; polymerase chain reaction.
 XX
 OS Synthetic.
 XX
 PN WO9527067-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 30-MAR-1995; 95MO-US04080.
 XX
 PR 24-FEB-1995; 95US-0394232.
 PR 01-APR-1994; 94US-0222612.
 PR 16-SEP-1994; 94US-0307410.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI Riggs MG, Sivaram M, Tudor SD;
 DR WPI: 1995-373510/48.
 DR N-PSDB; AAT04804.
 XX
 PT Bacillus steartothermophilus DNA encoding thermostable polymerase -
 PT for nucleic acid sequencing and amplification
 XX
 PS Claim 9; Page 92-94; 135pp; English.
 XX
 CC This is the amino acid sequence of the modified B.steartothermophilus
 CC thermostable DNA polymerase encoded by the clone Bst4. This clone
 CC encodes the sequence of a natural degradation product obtained when the
 CC DNA polymerase clone Bst3 (AAT04803) is expressed in E.coli. The same
 CC sequence is found when the full length clone (AAT04801) is expressed and
 CC cleaved with subtilisin to produce a "Klenow-type" DNA polymerase. This
 CC sequence starts with the Val residue at pos. 287 of the full length
 CC protein. This contrasts with a commercially available
 CC B.steartothermophilus DNA polymerase subtilisin fragment where the
 CC protein sequence starts with the Ala residue corresp. to pos. 290. The
 CC modified DNA polymerase can be used in primer extension reactions such as
 CC DNA sequencing or polymerase chain reaction.
 XX
 SQ Sequence 590 AA;
 Query Match 98.0%; Score 48; DB 16; Length 590;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAONLNTXKRE 11
 DB 435 LAONLNTXKRE 445
 RESULT 14
 ID AAE11057 standard; Protein; 592 AA.
 XX
 AC AAE11057;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Thermoclinomyces vulgaris Tvu DNA polymerase M285 truncated mutant.
 XX
 KW Reverse transcription; DNA polymerase; Thermoclinomyces vulgaris; Tvu;
 KW enzyme; amplification; hybridisation; mutant; mutin.
 XX
 OS Thermoclinomyces vulgaris.
 XX
 PN WO200164954-A1.
 XX
 PD 07-SEP-2001.

XX
 PF 28-FEB-2001; 2001MO-US06571.
 XX
 PR 02-MAR-2000; 2000US-0517871.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Gu T, Huang F, Hartnett JR;
 DR WPI: 2001-589872/66.
 DR N-PSDB; AAD18444.
 XX
 PT Reverse transcription of RNA using Tvu and Bst DNA polymerase,
 PT comprises combining the polymerase, RNA containing sample, a primer and
 PT reaction buffer having magnesium ions and reacting the mixture to
 obtain cDNA
 XX
 PS Claim 7; Fig 4; 118pp; English.
 XX
 CC The invention relates to reverse transcription of RNA templates using
 CC thermostable DNA polymerase such as Thermoclinomyces vulgaris (Tvu) and
 CC Bacillus steartothermophilus (Bst) DNA polymerase. Tvu and Bst DNA
 CC polymerase are useful for reverse transcription reactions which allows
 CC repetitive heating/cooling cycles without the requirement of fresh enzyme
 CC at each cooling step. The invention contemplates single-reaction RT-PCR
 CC wherein reverse transcription and amplification are performed in a single
 CC continuous procedure. Primers, template, nucleic acid triphosphates,
 CC and polymerase are used in the PCR process, which involves denaturation
 CC of target DNA, hybridisation of primers and synthesis of complementary
 CC strands. The present sequence is Thermoclinomyces vulgaris Tvu DNA
 CC polymerase M285 truncated mutant. This Tvu M285 mutant is obtained by
 CC deleting 284 amino acids from the N-terminal end of wild-type T. vulgaris
 CC Tvu DNA polymerase.
 XX
 SQ Sequence 592 AA;
 Query Match 98.0%; Score 48; DB 22; Length 592;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAONLNTXKRE 11
 DB 437 LAONLNTXKRE 447
 RESULT 15
 ID AAB85950 standard; Protein; 592 AA.
 XX
 AC AAB85950;
 XX
 DT 30-NOV-2001 (first entry)
 XX
 DE Tvu DNA polymerase deletion mutant M285.
 XX
 KW Thermostable; DNA polymerase; hyperthermophilic; DNA synthesis;
 KW Tvu DNA polymerase; nucleic acid amplification; reverse transcription;
 KW deletion mutant; M285.
 XX
 OS Thermoclinomyces vulgaris.
 XX
 PN WO200164838-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001MO-US06436.
 XX
 PR 02-MAR-2000; 2000US-0517439.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Gu T, Huang F, Hartnett JR;

```

XX DR WPI: 2001-557768/62.
XX DR N-PSDB: AAH47392.
XX PT Novel thermostable DNA polymerase derived from Thermactinomycetes
XX PT vulgaris for us in many recombinant DNA techniques, including
XX PT amplification, reverse transcription and sequencing reactions
PS Claim 13; Fig 6; 113pp; English.
XX
CC CC The invention provides compositions comprising thermostable DNA
CC CC polymerases derived from hyperthermophilic eubacteria, in particular,
CC CC a purified and isolated Thermactinomycetes vulgaris (Tvu) DNA polymerase
CC CC having DNA synthesis activity. It is useful for determining the
CC CC nucleotide base sequence of a DNA molecule. Tvu DNA polymerase can be
CC CC used for processes of high temperature nucleic acid amplification and
CC CC sequencing without substantial loss of DNA synthesis activity. The enzyme
CC CC can be used to perform high temperature reverse transcription in the
CC CC absence of manganese ions and in high throughput robotically-manipulated
CC CC procedures because greater enzymatic stability is retained at room
CC CC temperature. Using the Tvu DNA polymerase, reverse transcription
CC CC reactions can be conducted at increased temperatures so that secondary
CC CC structure is removed or limited. The present sequence represents the
CC CC amino acid sequence of a 5'-3' exonuclease deletion mutant form of
CC CC Tvu DNA polymerase called M285.
XX
SQ Sequence 592 AA;

```

```

Query Match 98.0%; Score 48; DB 22; Length 592;
Best Local Similarity 90.9%; Pred. No. 0.2;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LAQNLNIXRKE 11
   |||||
Db 437 LAQNLNITRKE 447

```

Search completed: January 15, 2003, 11:20:24
 Job time : 31.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 (Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec

Title: US-09-823-649A-7

Perfect score: 49

Sequence: 1 LAONLINRKE 11

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: Issued_Patents_AA:*
2: /cgn2_6/prodata/1/iaa/5A.COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5B.COMB.pep:*
4: /cgn2_6/prodata/1/iaa/5A.COMB.pep:*
5: /cgn2_6/prodata/1/iaa/5B.COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	48	98.0	294	3	US-09-057-969-7
2	48	98.0	378	3	US-09-057-969-6
3	48	98.0	457	3	US-09-057-969-5
4	48	98.0	539	2	US-09-057-969-4
5	48	98.0	587	2	US-08-642-684-2
6	48	98.0	587	2	US-08-436-664-23
7	48	98.0	587	3	US-09-135-642-23
8	48	98.0	587	3	US-08-394-232A-23
9	48	98.0	587	4	US-09-157-397-2
10	48	98.0	587	4	US-09-517-871-22
11	48	98.0	587	5	PCT-US95-04080-23
12	48	98.0	588	2	US-08-436-664-25
13	48	98.0	588	3	US-09-135-642-25
14	48	98.0	588	3	US-08-394-232A-25
15	48	98.0	588	4	US-09-157-397-4
16	48	98.0	588	4	US-09-517-871-6
17	48	98.0	588	5	PCT-US95-04080-25
18	48	98.0	589	2	US-08-642-684-4
19	48	98.0	590	2	US-08-436-664-27
20	48	98.0	590	3	US-09-135-642-27
21	48	98.0	590	3	US-08-394-232A-27
22	48	98.0	590	5	PCT-US95-04080-27
23	48	98.0	592	4	US-09-517-871-4
24	48	98.0	593	1	US-08-208-036-17
25	48	98.0	593	1	US-08-428-823-17
26	48	98.0	624	3	US-09-057-969-3
27	48	98.0	876	2	US-08-510-215A-2

28	48	98.0	876	2	US-08-633-476-2	Sequence 2, Appl
29	48	98.0	876	2	US-08-436-664-20	Sequence 20, Appl
30	48	98.0	876	2	US-08-436-664-32	Sequence 32, Appl
31	48	98.0	876	2	US-08-436-664-34	Sequence 34, Appl
32	48	98.0	876	3	US-09-135-642-32	Sequence 20, Appl
33	48	98.0	876	3	US-09-135-642-32	Sequence 32, Appl
34	48	98.0	876	3	US-09-135-642-34	Sequence 34, Appl
35	48	98.0	876	3	US-08-394-232A-20	Sequence 20, Appl
36	48	98.0	876	3	US-08-394-232A-32	Sequence 32, Appl
37	48	98.0	876	3	US-08-394-232A-34	Sequence 34, Appl
38	48	98.0	876	4	US-09-517-871-2	Sequence 2, Appl
39	48	98.0	876	4	US-09-517-871-20	Sequence 20, Appl
40	48	98.0	876	5	PCT-US95-04080-20	Sequence 20, Appl
41	48	98.0	876	5	PCT-US95-04080-32	Sequence 32, Appl
42	48	98.0	876	5	PCT-US95-04080-34	Sequence 34, Appl
43	48	98.0	877	1	US-08-208-036-14	Sequence 14, Appl
44	48	98.0	877	1	US-08-428-823-14	Sequence 14, Appl
45	48	98.0	954	3	US-09-057-969-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-057-969-7
Sequence 7, Application US/09057969
Patent No. 6013451
GENERAL INFORMATION:
APPLICANT: KONG, VICTOR THAI WONG
APPLICANT: PHANG, SENG MENG
APPLICANT: TAN, TIEN CHEE
TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA
TITLE OF INVENTION: POLYMERASE I (KLENOV) CLONES INCLUDING THOSE REDUCED TO
TITLE OF INVENTION: 3'-TO-5' EXONUCLEASE ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
STREET: 745 FIFTH AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10151
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,969
FILING DATE: 09-APR-1998
CLASSIFICATION: 4.35
ATTORNEY/AGENT INFORMATION:
NAME: KOMALSKI, THOMAS J.
REGISTRATION NUMBER: 32,147
REFERENCE/DOCKET NUMBER: 674507-2001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 294 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-09-057-969-7
Query Match: 98.0%; Score 48; DB 3; Length 294;
Best Local Similarity: 90.9%; Pred. No. 0.024;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
LAONLINRKE 11
|||||||

Db 109 LAONLITRKE 119

RESULT 2

US-09-057-969-6

Sequence 6, Application US/09057969

Patent No. 6013451

GENERAL INFORMATION:

APPLICANT: WONG, VICTOR THI WONG

APPLICANT: PHANG, SENG MENG

APPLICANT: TAN, TIEN CHE

TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA

TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FROMMER LAWRENCE & HAUG LLP

STREET: 745 FIFTH AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10151

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/057,969

FILING DATE: 09-APR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOMALSKI, THOMAS J.

REGISTRATION NUMBER: 32,147

REFERENCE/DOCKET NUMBER: 674507-2001

TELEPHONE: 212-588-0800

TELEFAX: 212-588-0500

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 378 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: amino acid

US-09-057-969-6

Query Match 98.0%; Score 48; DB 3; Length 378;
Best Local Similarity 90.9%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLITRKE 11

Db 193 LAONLITRKE 203

RESULT 3

US-09-057-969-5

Sequence 5, Application US/09057969

Patent No. 6013451

GENERAL INFORMATION:

APPLICANT: WONG, VICTOR THI WONG

APPLICANT: PHANG, SENG MENG

APPLICANT: TAN, TIEN CHE

TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA

TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FROMMER LAWRENCE & HAUG LLP

STREET: 745 FIFTH AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10151

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/057,969

FILING DATE: 09-APR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOMALSKI, THOMAS J.

REGISTRATION NUMBER: 32,147

REFERENCE/DOCKET NUMBER: 674507-2001

TELEPHONE: 212-588-0800

TELEFAX: 212-588-0500

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 457 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: amino acid

US-09-057-969-5

Query Match 98.0%; Score 48; DB 3; Length 457;
Best Local Similarity 90.9%; Pred. No. 0.038;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLITRKE 11

Db 272 LAONLITRKE 282

RESULT 4

US-09-057-969-4

Sequence 4, Application US/09057969

Patent No. 6013451

GENERAL INFORMATION:

APPLICANT: WONG, VICTOR THI WONG

APPLICANT: PHANG, SENG MENG

APPLICANT: TAN, TIEN CHE

TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA

TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: FROMMER LAWRENCE & HAUG LLP

STREET: 745 FIFTH AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10151

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/057,969

FILING DATE: 09-APR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: KOMALSKI, THOMAS J.

REGISTRATION NUMBER: 32,147

REFERENCE/DOCKET NUMBER: 674507-2001

TELEPHONE: 212-588-0800

TELEFAX: 212-588-0500

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: amino acid
US-09-057-969-1

Query Match
Best Local Similarity 90.9%; Score 48; DB 3; Length 539;
Pred. No. 0.045;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNTXKE 11
DB 354 LAONLNTXKE 364

RESULT 5
US-08-642-684-2
Sequence 2, Application US/08642684
Patent No. 5634253

GENERAL INFORMATION:
APPLICANT: HONG, GDO FAN
APPLICANT: FENG, ZHAI
APPLICANT: HUANG, WEI-HUA
TITLE OF INVENTION: A NEW DNA POLYMERASE WITH PROOF-READING
TITLE OF INVENTION: 3'-5' EXONUCLEASE ACTIVITY
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DAREY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,684
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: CHAPIN, MARLAINA K.
REGISTRATION NUMBER: 35,843
REFERENCE/DOCKET NUMBER: 4694/219502
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3711
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-642-684-2

Query Match
Best Local Similarity 98.0%; Score 48; DB 2; Length 587;
Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNTXKE 11
DB 432 LAONLNTXKE 442

RESULT 6
US-08-436-664-23
Sequence 23, Application US/08436664
Patent No. 5874282

GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARRA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
TITLE OF INVENTION: STEAROTHERMOPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,664
FILING DATE: 08-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/394,232
FILING DATE:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994

ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A.
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-436-664-23

Query Match
Best Local Similarity 98.0%; Score 48; DB 2; Length 587;
Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNTXKE 11
DB 432 LAONLNTXKE 442

RESULT 7
US-09-135-642-23
Sequence 23, Application US/09135642
Patent No. 6066483

GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARRA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
TITLE OF INVENTION: STEAROTHERMOPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/394,232A
3 FILING DATE:
4 CLASSIFICATION: 435
5 PRIOR APPLICATION DATA:
6 APPLICATION NUMBER: 08/307,410
7 FILING DATE: 16-SEP-1994
8 APPLICATION NUMBER: 08/222,612
9 FILING DATE: 16-SEP-1994
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Fisher, Carlos A
12 REGISTRATION NUMBER: 36,510
13 REFERENCE/DOCKET NUMBER: GP94003.CP2
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 619-535-2807
16 TELEFAX: 619-546-7929
17
18 INFORMATION FOR SEQ ID NO: 23:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 587 amino acids
21 TYPE: amino acid
22 STRANDEDNESS: single
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25 HYPOTHEICAL: NO
26 ANTI-SENSE: NO
27 FRAGMENT TYPE: internal
28 ORIGINAL SOURCE:
29 US-08-394-232A-23
30
31 Query Match 98.0%; Score 48; DB 3; Length 587;
32 Best Local Similarity 90.9%; Pred. No. 0.049;
33 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
34
35 OY 1 LAQNINIXRKE 11
36 ||||||| |||
37 Db 432 LAQNINITRKE 442
38
39 RESULT 9
40 US-09-157-397-2
41 Sequence 2, Application US/09157397
42 Patent No. 6165765
43 GENERAL INFORMATION:
44 APPLICANT: HONG, GUOFAN
45 TITLE OF INVENTION: DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE
46 TITLE OF INVENTION: SELECTIVE DISCRIMINATION AGAINST FLUORESCENT
47 TITLE OF INVENTION: DYE LABELED DEXOXYNUCLEOTIDES
48 FILE REFERENCE: homologues:listing
49 CURRENT APPLICATION NUMBER: US/09/157,397
50 CURRENT FILING DATE: 1998-09-21
51 EARLIER APPLICATION NUMBER: 08/544,643
52 EARLIER FILING DATE: 1995-10-18
53 EARLIER APPLICATION NUMBER: 08/642,684
54 EARLIER FILING DATE: 1996-05-03
55 NUMBER OF SEQ ID NOS: 11
56 SOFTWARE: PatentIn Ver. 2.0 - beta
57 SEQ ID NO 2
58 LENGTH: 587
59 TYPE: PRT
60 ORGANISM: Bacillus stearothermophilus
61 US-09-157-397-2
62
63 Query Match 98.0%; Score 48; DB 4; Length 587;
64 Best Local Similarity 90.9%; Pred. No. 0.049;
65 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
66
67 OY 1 LAQNINIXRKE 11
68 ||||||| |||
69 Db 432 LAQNINITRKE 442

```

RESULT 10
US-09-517-871-22
; Sequence 22, Application US/09517871
; Patent No. 6436677
; GENERAL INFORMATION:
; APPLICANT: Hartnett, John R.
; APPLICANT: Huang, Fen
; APPLICANT: Gu, Trent
; TITLE OF INVENTION: Method of Reverse Transcription
; FILE REFERENCE: PRMG-03833
; CURRENT APPLICATION NUMBER: US/09/517,871
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 22
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Bacillus stearotherophilus
US-09-517-871-22
Query Match 98.0%; Score 48; DB 4; Length 587;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LAONLNTXKE 11
|||||
DB 432 LAONLNTXKE 442
RESULT 11
PCT-US95-04080-23
; Sequence 23, Application PC/US9504080
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; TITLE OF INVENTION: STEAROTHEROPHILUS
; NUMBER OF SEQUENCES: 34
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04080
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-04080-23
Query Match 98.0%; Score 48; DB 5; Length 587;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LAONLNTXKE 11
|||||
DB 432 LAONLNTXKE 442
RESULT 12
US-08-436-664-25
; Sequence 25, Application US/08436664
; Patent No. 5874282
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATTHUR
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS

; TITLE OF INVENTION: STEAROTHEROPHILUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,664
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: 08/222,612
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fisher, Carlos A.
; REGISTRATION NUMBER: 36,510
; REFERENCE/DOCKET NUMBER: GP94003.CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-2807
; TELEFAX: 619-546-7929
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; US-08-436-664-25
Query Match 98.0%; Score 48; DB 2; Length 588;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LAONLNTXKE 11
|||||
DB 433 LAONLNTXKE 443
RESULT 13
US-09-135-642-25
; Sequence 25, Application US/09135642
; Patent No. 6066483
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATTHUR
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; TITLE OF INVENTION: STEAROTHEROPHILUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/135,642
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/394,232
FILING DATE:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-09-135-642-25

Query Match 98.0%; Score 48; DB 3; Length 588;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNIYRKE 11
DB 433 LAONLNIYRKE 443

RESULT 14
US-08-394-232A-25
Sequence 25, Application US/08394232A
Patent No. 6100078
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHEROPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSER: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,232A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 588 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-394-232A-25

Query Match 98.0%; Score 48; DB 3; Length 588;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNIYRKE 11
DB 433 LAONLNIYRKE 443

RESULT 15
US-09-157-397-4
Sequence 4, Application US/09157397
Patent No. 6165765
GENERAL INFORMATION:
APPLICANT: HUANG, WEI-HUA
TITLE OF INVENTION: DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE
TITLE OF INVENTION: SELECTIVE DISCRIMINATION AGAINST FLUORESCENT
FILE REFERENCE: HONGSEQUENCING
CURRENT APPLICATION NUMBER: US/09/157,397
CURRENT FILING DATE: 1998-09-21
EARLIER APPLICATION NUMBER: 08/544,643
EARLIER FILING DATE: 1995-10-18
EARLIER APPLICATION NUMBER: 08/642,684
EARLIER FILING DATE: 1996-05-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0 - beta
SEQ ID NO 4
LENGTH: 588
TYPE: PRT
ORGANISM: Bacillus stearotherophilus
US-09-157-397-4

Query Match 98.0%; Score 48; DB 4; Length 588;
Best Local Similarity 90.9%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LAONLNIYRKE 11
DB 433 LAONLNIYRKE 443

Search completed: January 15, 2003, 12:38:26
Job time: 10.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 12:36:21 ; Search time 26 Seconds

(without alignments)
129.748 Million cell updates/sec

Title: us-09-823-649a-7
Perfect score: 49
Sequence: 1 LAQNLNIXRKE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODE=frame+ p2n.model -DEV=11h
-O/cgn2_1/USPT.spool/US09823649/runat.14012003.151002.29127/app.query.fasta.1.1393
-DB=Issued Patents NA -OPMT=fastap -SUPFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09823649 -CGCN_1.1.35 -runat.14012003.151002.29127 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	98.0	890	3	US-09-057-869-12
2	48	98.0	1442	3	US-09-057-869-11
3	48	98.0	1379	3	US-09-057-869-10
4	48	98.0	1625	3	US-09-057-869-9
5	48	98.0	1764	2	US-08-642-684-1
6	48	98.0	1764	2	US-08-436-664-22
7	48	98.0	1764	3	US-09-135-642-22
8	48	98.0	1764	3	US-08-394-332A-22
9	48	98.0	1764	4	US-08-157-397-1
10	48	98.0	1764	4	US-09-157-397-3
11	48	98.0	1764	4	US-09-517-871-21
12	48	98.0	1764	5	PCT-US95-04080-22

13	48	98.0	1767	2	US-08-436-664-24	Sequence 24, Appl
14	48	98.0	1767	3	US-09-135-664-24	Sequence 24, Appl
15	48	98.0	1767	3	US-08-394-232A-24	Sequence 24, Appl
16	48	98.0	1767	4	US-09-517-871-5	Sequence 5, Appl1
17	48	98.0	1767	5	PCT-US95-04080-24	Sequence 24, Appl
18	48	98.0	1770	2	US-08-642-684-3	Sequence 3, Appl1
19	48	98.0	1773	2	US-08-436-664-26	Sequence 26, Appl
20	48	98.0	1773	3	US-09-135-664-26	Sequence 26, Appl
21	48	98.0	1773	3	US-08-394-232A-26	Sequence 26, Appl
22	48	98.0	1773	5	PCT-US95-04080-26	Sequence 26, Appl
23	48	98.0	1779	1	US-08-208-036-13	Sequence 13, Appl
24	48	98.0	1779	1	US-08-428-823-16	Sequence 16, Appl
25	48	98.0	1779	4	US-08-428-823-16	Sequence 16, Appl
26	48	98.0	1880	3	US-09-057-869-8	Sequence 3, Appl1
27	48	98.0	2631	1	US-08-208-036-13	Sequence 3, Appl1
28	48	98.0	2631	1	US-08-428-823-13	Sequence 13, Appl
29	48	98.0	2631	1	US-08-510-215A-1	Sequence 1, Appl1
30	48	98.0	2631	2	US-08-436-664-21	Sequence 21, Appl
31	48	98.0	2631	2	US-08-436-664-31	Sequence 31, Appl
32	48	98.0	2631	2	US-08-436-664-31	Sequence 31, Appl
33	48	98.0	2631	3	US-09-135-642-21	Sequence 33, Appl
34	48	98.0	2631	3	US-09-135-642-21	Sequence 33, Appl
35	48	98.0	2631	3	US-09-135-642-31	Sequence 31, Appl
36	48	98.0	2631	3	US-09-135-642-33	Sequence 33, Appl
37	48	98.0	2631	3	US-08-394-232A-21	Sequence 21, Appl
38	48	98.0	2631	3	US-08-394-232A-31	Sequence 31, Appl
39	48	98.0	2631	3	US-08-394-232A-33	Sequence 33, Appl
40	48	98.0	2631	4	US-09-517-871-1	Sequence 1, Appl1
41	48	98.0	2631	4	US-09-517-871-19	Sequence 19, Appl
42	48	98.0	2631	5	PCT-US95-04080-31	Sequence 31, Appl
43	48	98.0	2631	5	PCT-US95-04080-33	Sequence 33, Appl
44	48	98.0	2761	2	US-08-436-664-19	Sequence 19, Appl
45	48	98.0	2761	3	US-09-135-664-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-057-969-12
; Sequence 12, Application US/09057969
; Patent No. 6013451
; GENERAL INFORMATION:
; APPLICANT: WONG, VICTOR THAI WONG
; APPLICANT: PHANG, SENG MENG
; APPLICANT: TAN, TIEN CHYE
; TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA
; TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FROMER LAWRENCE & HAUG LLP
; STREET: 745 FIFTH AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057, 969
; FILING DATE: 09-APR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOMALSKI, THOMAS J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 674507-2001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:
LENGTH: 890 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-057-969-12

Alignment Scores:
Pred. No.: 0.0171 Length: 890
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-057-969-12 (1-890)

QY 1 LeuAglInasnlEuAsnle**Arglysglu 11
|||||
DB 325 CTGGCGCAAACTTGAACATTTACGGCAAGAA 357

RESULT 2

US-09-057-969-11
Sequence 11, Application US/09057969
Patent No. 6013451
GENERAL INFORMATION:
APPLICANT: WONG, VICTOR THI WONG
APPLICANT: PHANG, SENG MENG
APPLICANT: TAN, TIEN CHYE
TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA
TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO
NUMBER OF SEQUENCES: 3'-TO-5' EXONUCLEASE ACTIVITY
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,969
FILING DATE: 09-APR-1998
CORRESPONDENCE ADDRESS:
ADDRESS: FROMMER LAWRENCE & HAUG LLP
STREET: 745 FIFTH AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10151
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,969
FILING DATE: 09-APR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOMALSKI, THOMAS J.
REGISTRATION NUMBER: 32,147
REFERENCE/DOCKET NUMBER: 674507-2001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1142 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-057-969-11

Alignment Scores:
Pred. No.: 0.0233 Length: 1142
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-057-969-11 (1-1142)

QY 1 LeuAglInasnlEuAsnle**Arglysglu 11
|||||
DB 577 CTGGCGCAAACTTGAACATTTACGGCAAGAA 609

RESULT 3

US-09-057-969-10
Sequence 10, Application US/09057969
Patent No. 6013451
GENERAL INFORMATION:
APPLICANT: WONG, VICTOR THI WONG
APPLICANT: PHANG, SENG MENG
APPLICANT: TAN, TIEN CHYE
TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA
TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO
NUMBER OF SEQUENCES: 3'-TO-5' EXONUCLEASE ACTIVITY
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,969
FILING DATE: 09-APR-1998
CORRESPONDENCE ADDRESS:
ADDRESS: FROMMER LAWRENCE & HAUG LLP
STREET: 745 FIFTH AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10151
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,969
FILING DATE: 09-APR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOMALSKI, THOMAS J.
REGISTRATION NUMBER: 32,147
REFERENCE/DOCKET NUMBER: 674507-2001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-588-0800
TELEFAX: 212-588-0500
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1379 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-057-969-10

Alignment Scores:
Pred. No.: 0.0295 Length: 1379
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-057-969-10 (1-1379)

QY 1 LeuAglInasnlEuAsnle**Arglysglu 11
|||||
DB 814 CTGGCGCAAACTTGAACATTTACGGCAAGAA 846

RESULT 4

US-09-057-969-9
Sequence 9, Application US/09057969
Patent No. 6013451
GENERAL INFORMATION:
APPLICANT: WONG, VICTOR THI WONG
APPLICANT: PHANG, SENG MENG
APPLICANT: TAN, TIEN CHYE

1 TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA
2 TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO
3 TITLE OF INVENTION: 3'-TO-5' EXONUCLEASE ACTIVITY
4 NUMBER OF SEQUENCES: 17
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: FROMER LAWRENCE & HAUG LLP
7 STREET: 745 FIFTH AVENUE
8 CITY: NEW YORK
9 STATE: NEW YORK
10 COUNTRY: USA
11 ZIP: 10151
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: Patentin Release #1.0, Version #1.30
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/09/057,969
20 FILING DATE: 09-APR-1998
21 CLASSIFICATION: 435
22 ATTORNEY/AGENT INFORMATION:
23 NAME: KOMALSKI, THOMAS J.
24 REGISTRATION NUMBER: 32,147
25 REFERENCE/DOCKET NUMBER: 674507-2001
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: 212-588-0800
28 TELEFAX: 212-588-0500
29 INFORMATION FOR SEQ. ID NO.: 9:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 1625 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 MOLECULE TYPE: DNA (genomic)
36
37 US-09-057-969-9
38
39 Alignment Scores:
40 Pred. No.: 0.0362 Length: 1625
41 Score: 48.00 Matches: 10
42 Percent Similarity: 90.91% Conservative: 0
43 Best local Similarity: 90.91% Mismatches: 1
44 Query Match: 97.96% Indels: 0
45 DB: 3 Gaps: 0
46
47 US-09-823-649A-7 (1-11) x US-09-057-969-9 (1-1625)
48
49 QY 1 LeuAlaGlnAsnLeuAsnIle**ArGlySglu 11
50 Db 1060 CTGGCCCAAACTTGACATTACGCCCAAGCA 1092
51
52 RESULT 5
53 US-08-642-684-1
54 Sequence 1, Application US/08642684
55 Patent No. 5834253
56 GENERAL INFORMATION:
57 APPLICANT: HONG, GUO FAN
58 APPLICANT: FENG, ZHAI
59 APPLICANT: HUANG, WEI-HUA
60 TITLE OF INVENTION: A NEW DNA POLYMERASE WITH PROOF-READING
61 TITLE OF INVENTION: 3'-5' EXONUCLEASE ACTIVITY
62 NUMBER OF SEQUENCES: 15
63 CORRESPONDENCE ADDRESS:
64 ADDRESSEE: CUSHMAN, DAREY & CUSHMAN
65 STREET: 1100 NEW YORK AVENUE, N.W.
66 CITY: WASHINGTON
67 STATE: D.C.
68 COUNTRY: USA
69 ZIP: 20005
70
71 COMPUTER READABLE FORM:
72 MEDIUM TYPE: Floppy disk
73 COMPUTER: IBM PC compatible
74 OPERATING SYSTEM: PC-DOS/MS-DOS
75 SOFTWARE: Patentin Release #1.0, Version #1.25

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/642,684
3 FILING DATE: 03-MAY-1996
4 CLASSIFICATION: 435
5 ATTORNEY/AGENT INFORMATION:
6 NAME: CHAPIN, MARIANA K.
7 REGISTRATION NUMBER: 35,843
8 REFERENCE/DOCKET NUMBER: 4694/219502
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 202-861-3711
11 TELEFAX: 202-822-0944
12 TELUX: 6714627 CUSH
13 INFORMATION FOR SMO. ID NO.: 1:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 1764 base pairs
16 TYPE: nucleic acid
17 STRANDEDNESS: single
18 TOPOLOGY: linear
19 MOLECULE TYPE: DNA (genomic)
20
21 US-08-642-684-1
22
23 Alignment Scores:
24 Pred. No.: 0.0401 Length: 1764
25 Score: 48.00 Matches: 10
26 Percent Similarity: 90.91% Conservative: 0
27 Best local Similarity: 90.91% Mismatches: 1
28 Query Match: 97.96% Indels: 0
29 DB: 2 Gaps: 0
30
31 US-09-823-649A-7 (1-11) x US-08-642-684-1 (1-1764)
32
33 QY 1 LeuAlaGlnAsnLeuAsnIle**ArGlySglu 11
34 Db 1294 TTGGGGCAAACTTGACATTACGCCCAAGCA 1326
35
36 RESULT 6
37 US-08-436-664-22
38 Sequence 22, Application US/08436664
39 Patent No. 5874282
40 GENERAL INFORMATION:
41 APPLICANT: RIGGS, MICHAEL G.
42 APPLICANT: SIVARAM, MATHEOR
43 APPLICANT: TUDOR, STAKIA D.
44 TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
45 TITLE OF INVENTION: STEAROTHERMOPHILUS
46 NUMBER OF SEQUENCES: 34
47 CORRESPONDENCE ADDRESS:
48 ADDRESSEE: Gen Probe, Incorporated
49 STREET: 9080 Campus Point Drive
50 CITY: San Diego
51 STATE: CA
52 COUNTRY: USA
53 ZIP: 92121
54
55 COMPUTER READABLE FORM:
56 MEDIUM TYPE: Diskette
57 COMPUTER: IBM compatible
58 OPERATING SYSTEM: DOS
59 SOFTWARE: FastSeq Version 1.5
60 CURRENT APPLICATION DATA:
61 APPLICATION NUMBER: US/08/436,664
62 FILING DATE: 08-MAY-1995
63 CLASSIFICATION: 536
64 PRIOR APPLICATION DATA:
65 APPLICATION NUMBER: 08/394,232
66 FILING DATE:
67 APPLICATION NUMBER: 08/307,410
68 FILING DATE: 16-SEP-1994
69 APPLICATION NUMBER: 08/222,612
70 FILING DATE: 16-SEP-1994
71 ATTORNEY/AGENT INFORMATION:
72 NAME: FISHER, CARLOS A.
73 REGISTRATION NUMBER: 36,510
74 REFERENCE/DOCKET NUMBER: GP94003.CP2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1764 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1..1761
OTHER INFORMATION:
US-08-436-664-22

Alignment Scores:
Pred. No.: 0.0001 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-7 (1-11) x US-08-436-664-22 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
DB 1294 CTGGCCGAACACTTGACATTACGCCGAAGAA 1326

RESULT 7
US-09-135-642-22
Sequence 22, Application US/09135642
Patent No. 6066483
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: STARAM, MATTHOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
TITLE OF INVENTION: STEAROTHERMOPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/135,642
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/394,232
FILING DATE:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1764 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1..1761
OTHER INFORMATION:
US-09-135-642-22

Alignment Scores:
Pred. No.: 0.0401 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0

US-09-823-649a-7 (1-11) x US-09-135-642-22 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
DB 1294 CTGGCCGAACACTTGACATTACGCCGAAGAA 1326

RESULT 8
US-08-394-232a-22
Sequence 22, Application US/08394232A
Patent No. 6100078
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: STARAM, MATTHOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
TITLE OF INVENTION: STEAROTHERMOPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,232A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929


```
TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1761
; OTHER INFORMATION:
US-08-394-232A-22

Alignment Scores:
Pred. No.: 0.0401 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-7 (1-11) x US-08-394-232A-22 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
Db 1294 TTGGCGCAAACTTGACACTTACGCCCAAGAA 1326

RESULT 9
US-09-157-397-1
; Sequence 1, Application US/09157397
; Patent No. 6165765
; GENERAL INFORMATION:
; APPLICANT: HONG, GUOFAN
; TITLE OF INVENTION: DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE
; TITLE OF INVENTION: SELECTIVE DISCRIMINATION AGAINST FLUORESCENT
; TITLE OF INVENTION: DYE-LABELLED DIDEOXYNUCLEOTIDES
; FILE REFERENCE: hongsequencelisting
; CURRENT APPLICATION NUMBER: US/09/157,397
; EARLIER FILING DATE: 1998-09-21
; EARLIER APPLICATION NUMBER: 08/544,643
; EARLIER FILING DATE: 1995-10-18
; EARLIER APPLICATION NUMBER: 08/642,684
; EARLIER FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 1
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-09-157-397-1

Alignment Scores:
Pred. No.: 0.0401 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-7 (1-11) x US-09-157-397-1 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
Db 1294 TTGGCGCAAACTTGACACTTACGCCCAAGAA 1326

RESULT 10
US-09-157-397-3
```

```
; Sequence 3, Application US/09157397
; Patent No. 6165765
; GENERAL INFORMATION:
; APPLICANT: HONG, GUOFAN
; APPLICANT: HUANG, WEI-HUA
; TITLE OF INVENTION: DNA POLYMERASE HAVING ABILITY TO REDUCE INNATE
; TITLE OF INVENTION: SELECTIVE DISCRIMINATION AGAINST FLUORESCENT
; TITLE OF INVENTION: DYE-LABELLED DIDEOXYNUCLEOTIDES
; FILE REFERENCE: hongsequencelisting
; CURRENT APPLICATION NUMBER: US/09/157,397
; EARLIER FILING DATE: 1998-09-21
; EARLIER APPLICATION NUMBER: 08/544,643
; EARLIER FILING DATE: 1995-10-18
; EARLIER APPLICATION NUMBER: 08/642,684
; EARLIER FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 3
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-09-157-397-3

Alignment Scores:
Pred. No.: 0.0401 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-7 (1-11) x US-09-157-397-3 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
Db 1297 TTCCGCCAAGACTTGACACTTACGCCCAAGAA 1329

RESULT 11
US-09-517-871-21
; Sequence 21, Application US/09517871
; Patent No. 6436677
; GENERAL INFORMATION:
; APPLICANT: Hartnett, John R.
; APPLICANT: Huang, Ren
; APPLICANT: Gu, Trent
; TITLE OF INVENTION: Method of Reverse Transcription
; FILE REFERENCE: PRMG-03833
; CURRENT APPLICATION NUMBER: US/09/517,871
; EARLIER FILING DATE: 2000-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Bacillus stearothermophilus
US-09-517-871-21

Alignment Scores:
Pred. No.: 0.0401 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 4 Gaps: 0

US-09-823-649A-7 (1-11) x US-09-517-871-21 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
Db 1294 TTGGCGCAAACTTGACACTTACGCCCAAGAA 1326

RESULT 12
PCT-US95-04080-22
```

; Sequence 22, Application PC/TUS9504080
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; TITLE OF INVENTION: STEAROTHERMOPHILUS
; NUMBER OF SEQUENCES: 34
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04080
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1761
; PCT-US95-04080-22

Alignment Scores:
Pred. No.: 0.0401 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 5 Gaps: 0

US-09-823-649a-7 (1-11) x PCT-US95-04080-22 (1-1764)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArgIysGlu 11
|||||
Db 1294 CTGGCGCAAACTTGACATTCGCGCAAGAA 1326

RESULT 13
US-08-436-664-24
; Sequence 24, Application US/08436664
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATHOOR
; APPLICANT: TUDOR, STARRA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; TITLE OF INVENTION: STEAROTHERMOPHILUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,664
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994

APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1767 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..1764
; OTHER INFORMATION:
US-08-436-664-24

Alignment Scores:
Pred. No.: 0.0402 Length: 1767
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-7 (1-11) x US-08-436-664-24 (1-1767)
QY 1 LeuAlaGlnAsnLeuAsnIle**ArgIysGlu 11
|||||
Db 1297 CTGGCGCAAACTTGACATTCGCGCAAGAA 1329

RESULT 14
US-09-135-642-24
; Sequence 24, Application US/09135642
; Patent No. 6066483
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATHOOR
; APPLICANT: TUDOR, STARRA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; TITLE OF INVENTION: STEAROTHERMOPHILUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994

APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1..1764
OTHER INFORMATION:
US-09-135-642-24
Alignment Scores:
Pred. No.: 0.0402 Length: 1767
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0
US-09-823-649a-7 (1-11) x US-09-135-642-24 (1-1767)
Qy 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
|||||
Db 1297 CTGGCGCAAAACCTTGACATATACGCGCAAGAA 1329
RESULT 15
US-08-394-232A-24
Sequence 24, Application US/08394232A
Patent No. 6100078
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MAHMOOD
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILUS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,232A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:

NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1..1764
OTHER INFORMATION:
US-08-394-232A-24
Alignment Scores:
Pred. No.: 0.0402 Length: 1767
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 3 Gaps: 0
US-09-823-649a-7 (1-11) x US-08-394-232A-24 (1-1767)
Qy 1 LeuAlaGlnAsnLeuAsnIle***ArgLysGlu 11
|||||
Db 1297 CTGGCGCAAAACCTTGACATATACGCGCAAGAA 1329
Search completed: January 15, 2003, 12:58:58
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 15, 2003, 11:21:44 ; Search time 130.143 Seconds

(without alignments)
190.345 Million cell updates/sec

Title: US-09-823-649a-7
Perfect score: 49
Sequence: 1 LAQNLINXKE 11

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODEL=frame+ .p2n.model -DEV=xlh
-O=/cgn2.1/USPRO.spool/US09823649/runat_14012003.151001.29119/app.query.fasta_1.1393
-DB=N.Geneseq.101002 -OPMT=fastlap -SUFFIX=ring -MINMATCH=0.1 -LOOPTC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -NAMES=blusum62 -TRANS=human40.cdt
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649_ECGN_1.1.0_@runat_14012003.151001.29119 -NCPU=6 -ICPU=3
-NO_XLPEXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	98.0	1763	18 AAT36360	Bst DNA polymerase
2	48	98.0	1764	18 AAT04802	B.stearothermophil
3	48	98.0	1764	18 AAT90819	Bacillus stearother
4	48	98.0	1764	21 AAA08544	B. stearothermophil
5	48	98.0	1764	21 AAA08545	B. stearothermophil
6	48	98.0	1764	22 AAD18459	B. stearothermophil
7	48	98.0	1767	16 AAT04803	B. stearothermophil
8	48	98.0	1767	22 AAD18445	Thermoclinomyces
9	48	98.0	1767	22 AA47393	Tvu DNA polymerase
10	48	98.0	1770	18 AAT90820	B. stearothermophil
11	48	98.0	1773	16 AAT04804	B. stearothermophil
12	48	98.0	1779	13 AAQ31651	Bacillus caldotena
13	48	98.0	1779	22 AAD18444	Thermoclinomyces
14	48	98.0	1779	22 AAH47392	Tvu DNA polymerase
15	48	98.0	2631	16 AAT04800	B. stearothermophil
16	48	98.0	2631	16 AAT04805	B. stearothermophil
17	48	98.0	2631	16 AAT04806	B. stearothermophil
18	48	98.0	2631	18 AAT75434	Bacillus caldotena
19	48	98.0	2631	18 AAT75432	DNA encoding Bacil
20	48	98.0	2631	18 AAT75433	DNA encoding Bacil
21	48	98.0	2631	22 AAD18443	Thermoclinomyces
22	48	98.0	2631	22 AAD18448	Bacillus stearothe
23	48	98.0	2631	22 AAH47391	Nucleotide sequenc
24	48	98.0	2761	16 AAT04801	B. stearothermophil
25	48	98.0	2969	19 AAV67204	DNA encoding a DNA
26	48	98.0	3246	14 AA054170	B. stearothermophil
27	48	98.0	3252	13 AA031650	Bacillus caldotena
28	48	98.0	3317	18 AAT97089	Bacillus stearothe
29	48	81.6	3904	23 ABT26758	Drosophila melanog
30	39	79.6	2083	24 ABG70553	Listeria monocytog
31	39	79.6	2543	24 ABK74233	Bacillus lichenifo
32	39	79.6	2545	19 AAV11723	B. pallidus DNA po
33	39	79.6	3255	19 AAV11724	B. pallidus DNA po
34	39	79.6	3831	19 AAV11725	B. pallidus DNA po
35	39	79.6	49526	24 ABQ67195	Listeria innocua c
36	37	77.6	349	21 AAC24266	Human secreted pro
37	37	77.6	3177	22 AAS13624	CDNA sequence enco
38	37	75.5	3591	23 AB117784	Drosophila melanog
39	37	75.5	3771	23 AB117768	Drosophila melanog
40	36	73.5	491	23 AAS89981	DNA encoding novel
41	36	73.5	2649	23 AAS52934	Enterococcus faeca
42	36	73.5	2787	23 AAS52883	E. coli DNA for ce
43	36	73.5	5060	18 AAT78171	Plasmid pKlen-TBD1
44	36	73.5	5062	18 AAT78170	Plasmid pSR-2 enco
45	36	73.5	8033	20 AAX13058	Enterococcus faeca

ALIGNMENTS

RESULT 1
AAT36360
AAT36360 standard: DNA, 1763 bp.

XX AAT36360;
DT 07-MAY-1997 (first entry)

XX Bst DNA polymerase large fragment DNA.

DE DNA polymerase; Bst Pol I; thermostable enzyme; exonuclease;

KM DNA sequencing; strand displacement amplification; ss.

XX Bacillus stearothermophilus.

XX EF757100-AL.

XX 05-FEB-1997.

```

XX 31-JUL-1996; 96EP-0202169.
PF 02-AUG-1995; 95US-0510215.
XX (NEME) NEW ENGLAND BIOLABS INC.
XX
XX Allotta JM, Kong H, Pelletier JT;
XX WPI; 1997-111048/11.
DR P-PSDB; AAM10597.
XX
XX DNA encoding recombinant thermophilic Bacillus stearothermophilus
PT DNA polymerase I - substantially free of 3'-5', and opt. also 5'-3',
PT exonuclease activity
XX
XX Claim 3; Page 12-16; 34pp; English.
XX
XX A DNA construct (AAT936360) codes for Bacillus stearothermophilus
CC (Bst) DNA polymerase large fragment (AAM10597) which lacks the
CC N-terminal 3'-5' exonuclease domain of the native enzyme. It
CC comprises nucleotides 868-2631 of the full-length Bst polymerase
CC coding sequence, and was ocl'd. by PCR cloning (see also AAT60826-27)
CC using Bst genomic DNA as template. The DNA construct can be used
CC to produce recombinant Bst polymerase large fragment, pref. as a
CC fusion protein with maltose binding protein, in transformed host
CC cells. The product can be used for nucleic acid sequencing and
CC strand displacement amplification.
XX
XX SQ Sequence 1763 BP; 474 A; 416 C; 486 G; 387 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.541* length: 1763
XX Score: 48.00 Matches: 10
XX Percent Similarity: 90.91% Conservative: 0
XX Best Local Similarity: 90.91% Mismatches: 1
XX Query Match: 97.96% Indels: 0
XX DB: 18 Gaps: 0
XX
XX US-09-823-649a-7 (1-11) x AAT936360 (1-1763)
XX
XX QY 1 LeuAlaGlnAsnLeuAsnIle**ArqLysGlu 11
XX |||||||
XX Db 1294 TTGGCGCAAACTTGACATTCAGCGCAAGAA 1326
XX
XX RESULT 2
XX ID AAT04802 standard; DNA; 1764 BP.
XX
XX AC AAT04802;
XX
XX DT 30-MAY-1996 (first entry)
XX
XX DE B. stearothermophilus DNA polymerase clone Bst2.
XX
XX KW Bacillus stearothermophilus; thermostable DNA polymerase; PCR; primer;
KW amplification; probe; Bacillus caldotenax; exonuclease; recombination;
KW sequencing; polymerase chain reaction; ss.
XX
XX OS Synthetic.
XX
XX PN W09527067-A1.
XX
XX PD 12-OCT-1995.
XX
XX PF 30-MAR-1995; 95MO-US04080.
XX
XX PR 24-FEB-1995; 95US-0394232.
XX PR 01-APR-1994; 94US-0222612.
XX PR 16-SEP-1994; 94US-0307410.
XX
XX PA (GENP-) GEN-PROBE INC.

```

```

PI Riggs MG, Sivaram M, Tudor SD;
XX
XX WPI; 1995-373510/48.
DR P-PSDB; AAR81039.
XX
XX Bacillus stearothermophilus DNA encoding thermostable polymerase
PT for nucleic acid sequencing and amplification
XX
XX Claim 9; Page 83-85; 135pp; English.
XX
XX This is the nucleotide sequence of the clone Bst2 encoding a modified
CC Bacillus stearothermophilus thermostable DNA polymerase which lacks the
CC complete 5'-3' exonuclease domain. The gene was constructed by ligating
CC the HindIII-SalI fragment from the clone pGemBst2.1ss1 contg. 3' end of
CC the gene with the StyI-HindIII fragment of the clone pGemBst5' end contg.
CC the 5' end of the DNA polymerase gene. The StyI-HindIII fragment covers
CC the mid-region of the gene up to and including the sequences encoding
CC the 3'-5' exonuclease but not the 5'-3' exonuclease domain. The
CC remainder of the 5' end of the gene was reconstituted by the ligated and
CC extended oligonucleotides AAT04816-17. The modified DNA polymerase can
CC be used in primer extension reactions such as DNA sequencing or
CC polymerase chain reaction.
XX
XX SQ Sequence 1764 BP; 430 A; 456 C; 534 G; 344 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.541 length: 1764
XX Score: 48.00 Matches: 10
XX Percent Similarity: 90.91% Conservative: 0
XX Best Local Similarity: 90.91% Mismatches: 1
XX Query Match: 97.96% Indels: 0
XX DB: 16 Gaps: 0
XX
XX US-09-823-649a-7 (1-11) x AAT04802 (1-1764)
XX
XX QY 1 LeuAlaGlnAsnLeuAsnIle**ArqLysGlu 11
XX |||||||
XX Db 1294 CTGGCGCAAACTTGACATTCAGCGCAAGAA 1326
XX
XX RESULT 3
XX ID AAT90819 standard; DNA; 1764 BP.
XX
XX AC AAT90819;
XX
XX DT 14-APR-1998 (first entry)
XX
XX DE Bacillus stearothermophilus DNA polymerase.
XX
XX KW Bacillus stearothermophilus; DNA polymerase; DNA amplification;
KW DNA replication; exonuclease; proof reading; ss.
XX
XX OS Bacillus stearothermophilus.
XX
XX FH Key Location/Qualifiers
XX FT CDS 1..1764
XX FT /tag= a
XX FT /note= "The start codon is not indicated in the
XX specification"
XX
XX JP09220087-A.
XX
XX PD 26-AUG-1997.
XX
XX PF 17-MAY-1996; 96JP-0160402.
XX
XX PR 03-MAY-1996; 96US-0642684.
XX PR 18-OCT-1995; 95US-0544643.
XX
XX PA (HONGG/) HONG G F.
XX PA (HUAN/) HUANG W.
XX PA (ZHAJ/) ZHAI F.
XX PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACADE.

```

```
XX
PI Hong GF, Huang W, Zhai F;
XX
DR WPI: 1997-474304/44.
DR P-PSDB: AAY30093.
XX
PT DNA polymerase from Bacillus stearothermophilus - used in DNA
PT amplification and sequencing methods having lower frequency of
PT mismatch
XX
PS Claim 4: Page 3: 32pp; Japanese.
XX
CC This is the DNA sequence of the Bacillus stearothermophilus strain DNA
CC polymerase. This heat stable DNA polymerase has a peptide sequence
CC of 587 amino acids. This DNA polymerase can be used in a method for the
CC replication of a DNA chain. It can also be used in a method for the
CC determination of the sequence of DNA chain. As the DNA polymerase has a
CC proof reading 3'-5' exonuclease activity, possible mismatches of base
CC pairing that occurs in current methods of DNA sequence determination of
CC a DNA clone can be avoided.
XX
SQ Sequence 1764 BP; 477 A; 415 C; 486 G; 386 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.541 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 18 Gaps: 0
XX
US-09-823-649A-7 (1-11) x AAT90819 (1-1764)
QY 1 leuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
ID |||||
DB 1294 TTGGCGCAAAACTTGACATTAACGCCCAAGAA 1326
XX
RESULT 4
AA08544 standard; DNA: 1764 BP.
XX
AC AAA08544;
XX
DT 19-JUL-2000 (first entry)
XX
DE B. stearothermophilus 320 modified DNA polymerase coding sequence.
XX
KM Modified DNA polymerase; DNA sequencing: 3' to 5' exonuclease activity;
KM random primer labeling; site-directed mutagenesis; ss.
XX
OS Synthetic.
XX
OS Bacillus stearothermophilus.
XX
PN WO200017330-A1.
XX
PD 30-MAR-2000.
XX
PF 14-JAN-1999; 99WO-1B00146.
XX
PR 21-SEP-1998; 98US-0157397.
XX
PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.
XX
PI Hong G, Huang W;
XX
DR WPI: 2000-283563/24.
DR P-PSDB: AAY91930.
XX
PT New DNA polymerase, useful particularly in DNA sequencing, incorporates
PT all four dideoxynucleotide triphosphate terminators with about equal
PT efficiency
XX
PS Claim 8; Page 93-95; 109pp; English.
```

```
XX
CC The present sequence encodes a modified Bacillus stearothermophilus
CC DNA polymerase (AAY91930) that, during DNA sequencing, effectively
CC incorporates all four fluorescent dye-labeled dideoxynucleotide
CC triphosphates (ddNTP), or their analogs, and reduces the discrimination,
CC observed with the unmodified enzyme, for incorporation of labeled
CC ddCTP and ddATP, is new. The modified DNA polymerase, has proofreading
CC 3'-5' exonuclease activity during DNA sequencing from a template and the
CC polymerase excises mismatched nucleotides from the 3' terminus at a
CC faster rate than it removes correctly matched nucleotides. The
CC polymerase is used particularly for DNA sequencing, also for filling in
CC 5'-overhangs, synthesis of DNA probes by random primer labeling and in
CC site-directed mutagenesis, (claimed).
XX
SQ Sequence 1764 BP; 476 A; 418 C; 485 G; 385 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.541 Length: 1764
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 97.96% Indels: 0
DB: 21 Gaps: 0
XX
US-09-823-649A-7 (1-11) x AA08544 (1-1764)
QY 1 leuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
ID |||||
DB 1297 TTGGCGCAAAACTTGACATTAACGCCCAAGAA 1329
XX
RESULT 5
AA08545 standard; DNA: 1764 BP.
XX
AC AA08545;
XX
DT 19-JUL-2000 (first entry)
XX
DE B. stearothermophilus 320 wild type DNA polymerase coding sequence.
XX
KM Modified DNA polymerase; DNA sequencing: 3' to 5' exonuclease activity;
KM random primer labeling; site-directed mutagenesis; wild type; ss.
XX
OS Bacillus stearothermophilus.
XX
PN WO200017330-A1.
XX
PD 30-MAR-2000.
XX
PF 14-JAN-1999; 99WO-1B00146.
XX
PR 21-SEP-1998; 98US-0157397.
XX
PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.
XX
PI Hong G, Huang W;
XX
DR WPI: 2000-283563/24.
DR P-PSDB: AAY91931.
XX
PT New DNA polymerase, useful particularly in DNA sequencing, incorporates
PT all four dideoxynucleotide triphosphate terminators with about equal
PT efficiency
XX
PS Example 2; Page 80-82; 109pp; English.
XX
CC The present sequence encodes a wild type Bacillus stearothermophilus
CC DNA polymerase (AAY91931) that is modified (see AAY91930) so that during
CC DNA sequencing, it effectively incorporates all four fluorescent
CC dye-labeled dideoxynucleotide triphosphates (ddNTP), or their analogs,
CC and reduces the discrimination, observed with the unmodified enzyme, for
CC incorporation of labeled ddCTP and ddATP, is new. The modified DNA
CC polymerase, has proofreading 3'-5' exonuclease activity during DNA
```

CC sequencing from a template and the polymerase excises mismatched
 CC nucleotides from the 3' terminus at a faster rate than it removes
 CC correctly matched nucleotides. The polymerase is used particularly for
 CC DNA sequencing, also for filling in 5'-overhangs, synthesis of DNA
 CC probes by random primer labeling and in site-directed mutagenesis,
 CC (claimed).

XX SQ Sequence 1764 BP; 477 A; 415 C; 486 G; 386 T; 0 other;

Alignment Scores:

Pred. No.:	0.541	Length:	1764
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	21	Gaps:	0

US-09-823-649A-7 (1-11) x AAA08545 (1-1764)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
 |||||
 Db 1294 TTGGCGCAAACTTGACATTACCGCAAGAA 1326

RESULT 6
 AAD18459
 ID AAD18459 standard; DNA: 1764 BP.
 AC AAD18459;
 AC AAD18459;
 DT 18-DEC-2001 (first entry)
 XX
 DE Bacillus stearothermophilus (Bst) DNA polymerase 1 fragment DNA.
 XX
 KW Reverse transcription; DNA polymerase 1; Bacillus stearothermophilus;
 KW Bst; enzyme; amplification; hybridisation; ds.
 XX
 OS Bacillus stearothermophilus.

Key	Location/Qualifiers
FT CDS	1..1764
FT	/tag= a
FT	/product= "Bacillus stearothermophilus DNA polymerase 1"
FT	/note= "CDS does not include start codon"
FT	/partial

XX PN WO200164954-A1.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06571.
 XX
 PR 02-MAR-2000; 2000US-0517871.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Gu T, Huang F, Hartnett JR;
 XX
 DR WPI: 2001-589872/66.
 DR P-PSDB; AAE11060.
 XX
 PT Reverse transcription of RNA using Tvu and Bst DNA polymerase,
 PT comprises combining the polymerase, RNA containing sample, a primer and
 PT reaction buffer having magnesium ions and reacting the mixture to
 PT obtain cDNA -
 XX
 XX Claim 9; Fig 9; 118pp; English.
 XX
 PS The invention relates to reverse transcription of RNA templates using
 CC thermostable DNA polymerase such as Thermococcus vulgarens (Tvu) and
 CC Bacillus stearothermophilus (Bst) DNA polymerase. Tvu and Bst DNA
 CC polymerase are useful for reverse transcription reactions which allows
 CC repetitive heating/cooling cycles without the requirement of fresh enzyme
 CC at each cooling step. The invention contemplates single-reaction RT-PCR

CC wherein reverse transcription and amplification are performed in a single
 CC continuous procedure. Primers, template, nucleoside triphosphates,
 CC appropriate reaction buffer having magnesium ions, reaction conditions,
 CC and polymerase are used in the PCR process, which involves denaturation
 CC of target DNA, hybridisation of primers and synthesis of complementary
 CC strands. The present DNA sequence is Bacillus stearothermophilus (Bst)
 CC DNA polymerase 1 fragment which lacks 5' to 3' exonuclease activity.

XX SQ Sequence 1764 BP; 477 A; 415 C; 486 G; 386 T; 0 other;

Alignment Scores:

Pred. No.:	0.541	Length:	1764
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-7 (1-11) x AAD18459 (1-1764)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
 |||||
 Db 1294 TTGGCGCAAACTTGACATTACCGCAAGAA 1326

RESULT 7
 AAT04803
 ID AAT04803 standard; DNA: 1767 BP.
 AC AAT04803;
 AC AAT04803;
 XX
 DT 30-MAY-1996 (first entry)
 XX
 DE B.stearothermophilus DNA polymerase clone Bst3.
 XX
 KW Bacillus stearothermophilus; thermostable DNA polymerase; PCR; primer;
 KW amplification; probe; Bacillus caldwellii; exonuclease; recombination;
 KW sequencing; polymerase chain reaction; ss.
 XX
 OS Synthetic.
 XX
 CS Synthetic.
 XX
 PN WO9527067-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 30-MAR-1995; 95WO-US04080.
 XX
 PR 24-FEB-1995; 95US-0394232.
 PR 01-APR-1994; 94US-0222612.
 PR 16-SEP-1994; 94US-0307410.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI Riggs MG, Sivaram M, Tudor SD;
 XX
 DR WPI: 1995-373510/48.
 DR P-PSDB; AAR80140.
 XX
 PT Bacillus stearothermophilus DNA encoding thermostable polymerase -
 PT for nucleic acid sequencing and amplification
 XX
 XX Claim 9; Page 87-90; 135pp; English.
 XX
 PS This is the nucleotide sequence of the clone Bst3 encoding a modified
 CC Bacillus stearothermophilus thermostable DNA polymerase which contains a
 CC deletion in the 5'-3' exonuclease domain. The gene was constructed by
 CC digestion and recombination of the full length gene (AAT04801) to remove
 CC a 195 bp HincII-PvuII fragment from the region encoding the 5'-3'
 CC exonuclease. The deleted fragment removes 65 amino acids (residues
 CC 178-242 of the protein) including 2 Gly residues though to correspond to
 CC 2 amino acids in the E.coli DNA polymerase I necessary for 5'-3'
 CC exonuclease activity. The modified DNA polymerase can be used in primer
 CC extension reactions such as DNA sequencing or polymerase chain reaction.

Tue Jan 21 14:46:22 2003

ns-00-823-649a-7 www

ns-00-823-649a-7

SQ Sequence 1767 BP; 431 A; 457 C; 535 G; 344 T; 0 other;

Alignment Scores:

Pred. No.:	0.542	Length:	1767
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	16	Gaps:	0

US-09-823-649A-7 (1-11) x AAT04803 (1-1767)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
 |||||
 DB 1297 CTGGCGCAAAACTTGAAACATTCACGCCCAAGAA 1329

RESULT 8

ID AAD18445 standard; DNA; 1767 BP.

AC AAD18445;

DT 18-DEC-2001 (first entry)

DE Thermocactinomycetes vulgaris DNA polymerase T289M truncated mutant DNA.

KW Reverse transcription; DNA polymerase; Thermocactinomycetes vulgaris; Tvu;

XX enzyme; amplification; hybridisation; mutant; muten; ds.

OS Thermocactinomycetes vulgaris.

XX Synthetic.

XX Key Location/Qualifiers

FT CDS 1..1767

FT /tag= a

FT /product= "Thermocactinomycetes vulgaris DNA polymerase

FT T289M truncated mutant"

FT /transl_except= (pos:1..3, aa:Thr)

FT mutation

FT replace (2, C)

XX /tag= b

XX MO200164954-A1.

PD 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US06571.

PR 02-MAR-2000; 2000US-0517871.

XX (PROM-) PROMEGA CORP.

PA (PROM-) PROMEGA CORP.

PI Gu T, Huang F, Hartnett JR;

XX WPI; 2001-589872/66.

DR P-PSDB; AAB11058.

XX Reverse transcription of RNA using Tvu and Bst DNA polymerase,

PT comprises combining the polymerase, RNA containing sample, a primer and

PT reaction buffer having magnesium ions and reacting the mixture to

PT obtain cDNA

XX Example 11; Fig 5; 118pp; English.

XX The invention relates to reverse transcription of RNA templates using

CC thermostable DNA polymerase such as Thermocactinomycetes vulgaris (Tvu) and

CC Bacillus stearothermophilus (Bst) DNA polymerase. Tvu and Bst DNA

CC polymerase are useful for reverse transcription reactions which allows

CC repetitive heating/cooling cycles without the requirement of fresh enzyme

CC at each cooling step. The invention contemplates single-reaction RT-PCR

CC wherein reverse transcription and amplification are performed in a single

CC continuous procedure. Primers, template, nucleoside triphosphates,

CC appropriate reaction buffer having magnesium ions, reaction conditions,

CC and polymerase are used in the PCR process, which involves denaturation

CC of target DNA, hybridisation of primers and synthesis of complementary
 CC strands. The present sequence is Thermocactinomycetes vulgaris (Tvu) DNA
 CC polymerase T289M truncated mutant DNA. This Tvu T289M mutant DNA is
 CC obtained by deleting 864 bases from the 5' end of wild-type Tvu DNA
 CC polymerase gene and replacing 'C' at position 866 with 'T' in the wild-
 CC type Tvu DNA polymerase gene.

SQ Sequence 1767 BP; 427 A; 457 C; 539 G; 344 T; 0 other;

Alignment Scores:

Pred. No.:	0.542	Length:	1767
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-7 (1-11) x AAD18445 (1-1767)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArgLysGlu 11
 |||||
 DB 1297 CTGGCGCAAAACTTGAAACATTCACGCCCAAGAA 1329

RESULT 9

ID AAH47393 standard; DNA; 1767 BP.

AC AAH47393;

DT 30-NOV-2001 (first entry)

DE Tvu DNA polymerase deletion mutant T289M encoding DNA.

KW Thermocactinomycetes vulgaris; hyperthermophilic; DNA synthesis;

KW Tvu DNA polymerase; nucleic acid amplification; reverse transcription;

KW deletion mutant; T289M; ds.

XX Thermocactinomycetes vulgaris.

XX Key Location/Qualifiers

FT CDS 1..1767

FT /tag= a

FT /product= "Tvu DNA polymerase M285"

XX MO200164838-A2.

PD 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US06436.

PR 02-MAR-2000; 2000US-0517439.

XX (PROM-) PROMEGA CORP.

PA (PROM-) PROMEGA CORP.

PI Gu T, Huang F, Hartnett JR;

XX WPI; 2001-557768/62.

DR P-PSDB; AAB85951.

XX Novel thermostable DNA polymerase derived from Thermocactinomycetes

PT vulgaris for use in many recombinant DNA techniques, including

PT amplification, reverse transcription and sequencing reactions

XX Claim 17; Fig 7; 113pp; English.

XX The invention provides compositions comprising thermostable DNA

CC polymerases derived from hyperthermophilic eubacteria, in particular,

CC a purified and isolated Thermocactinomycetes vulgaris (Tvu) DNA polymerase

CC having DNA synthesis activity. It is useful for determining the

CC nucleotide base sequence of a DNA molecule. Tvu DNA polymerase can be

CC used for processes of high temperature nucleic acid amplification and

CC sequencing without substantial loss of DNA synthesis activity. The enzyme

CC can be used to perform high temperature reverse transcription in the

CC absence of manganese ions and in high throughput robotically-manipulated
 CC procedures because greater enzymatic stability is retained at room
 CC temperature. Using the Tvu DNA polymerase, reverse transcription
 CC reactions can be conducted at increased temperatures so that secondary
 CC structure is removed or limited. The present sequence represents the
 CC nucleotide sequence encoding a 5'-3' exonuclease deletion mutant form of
 CC Tvu DNA polymerase called T289M.
 XX
 SQ Sequence 1767 BP; 427 A; 457 C; 539 G; 344 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.542 Length: 1767
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 DB: 22 Gaps: 0
 US-09-823-649a-7 (1-11) x AAH47393 (1-1767)
 QY 1 LeuAAGlnAsnLeuAsnIle**ArgLySGlu 11
 DB 1297 CTGGCGCAAACTTGACATTACGCCCAAGAA 1329
 RESULT 10
 AAT90820
 ID AAT90820 standard; DNA; 1770 BP.
 XX
 AC AAT90820;
 XX
 DT 14-Apr-1998 (first entry)
 XX
 DE B. steaerothermophilus.BM4 polymerase encoding full length DNA sequence.
 XX
 KW Bacillus steaerothermophilus; DNA polymerase; PCR; primer;
 XX DNA replication; exonuclease; proof reading; ss.
 XX
 OS Bacillus steaerothermophilus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1770
 FT /tag= a
 PN JP09220087-A.
 XX
 PD 26-AUG-1997.
 XX
 PF 17-MAY-1996; 96JP-0160402.
 XX
 PR 03-MAY-1996; 96US-0642684.
 PR 18-OCT-1995; 95US-0544643.
 XX
 PA (HONGG/) HONG G. F.
 PA (HUAN/) HUANG W.
 PA (ZHAI/) ZHAI F.
 PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.
 XX
 PI Hong GF, Huang W, Zhai F;
 XX
 DR WPI: 1997-474304/44.
 DR P-PSDB: AAM30094.
 XX
 PT DNA polymerase from Bacillus steaerothermophilus - used in DNA
 PT amplification and sequencing methods having lower frequency of
 PT mismatch
 XX
 PS Claim 13: Pages 27-28; 32pp; Japanese.
 CC This DNA sequence of the Bacillus steaerothermophilus strain encodes a
 CC heat stable DNA polymerase. This DNA polymerase has a peptide sequence
 CC of 589 amino acids. This DNA polymerase can be used in a method for the
 CC replication of a DNA chain. It can also be used in a method for the
 CC determination of the sequence of DNA chain. As the DNA polymerase has a

CC proof reading 3'-5' exonuclease activity, possible mismatches of base
 CC pairing that occurs in current methods of DNA sequence determination of
 CC a DNA clone can be avoided.
 XX
 SQ Sequence 1770 BP; 480 A; 415 C; 488 G; 387 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.543 Length: 1770
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 97.96% Indels: 0
 DB: 18 Gaps: 0
 US-09-823-649a-7 (1-11) x AAT90820 (1-1770)
 QY 1 LeuAAGlnAsnLeuAsnIle**ArgLySGlu 11
 DB 1300 TTGGCGCAAACTTGACATTACGCCCAAGAA 1332
 RESULT 11
 AAT04804
 ID AAT04804 standard; DNA; 1773 BP.
 XX
 AC AAT04804;
 XX
 DT 30-MAY-1996 (first entry)
 XX
 DE B.steaerothermophilus DNA polymerase clone Bst4.
 XX
 KW Bacillus steaerothermophilus; thermostable DNA polymerase; PCR; primer;
 KW amplification; probe; Bacillus caldotenax; exonuclease; recombination;
 KW sequencing; polymerase chain reaction; ss.
 XX
 OS Synthetic.
 XX
 PN W09527067-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 30-MAR-1995; 95WO-US04080.
 XX
 PR 24-FEB-1995; 95US-0394232.
 PR 01-APR-1994; 94US-0222612.
 PR 16-SEP-1994; 94US-0307410.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI Riggs MG, Sivaram M, Tudor SD;
 XX
 DR WPI: 1995-374510/48.
 DR P-PSDB: AAR0141.
 XX
 PT Bacillus steaerothermophilus DNA encoding thermostable polymerase -
 PT for nucleic acid sequencing and amplification
 XX
 PS Claim 9: Page 92-94; 135pp; English.
 CC This is the nucleotide sequence of the clone Bst4 encoding a modified
 CC Bacillus steaerothermophilus thermostable DNA polymerase. This clone
 CC encodes the sequence of a natural degradation product obtained when the
 CC DNA polymerase clone Bst3 (AAT04803) is expressed in E.coli. The same
 CC sequence is found when the full length clone (AAT04801) is expressed and
 CC cleaved with subtilisin to produce a "Klenow-type" DNA polymerase. This
 CC sequence starts with the Val residue at pos. 287 of the full length
 CC protein. This contrasts with a commercially available
 CC B.steaerothermophilus DNA polymerase subtilisin fragment where the
 CC protein sequence starts with the Ala residue corresp. to pos. 290. The
 CC modified DNA polymerase can be used in primer extension reactions such as
 CC DNA sequencing or polymerase chain reaction.
 XX
 SQ Sequence 1773 BP; 433 A; 460 C; 535 G; 345 T; 0 other;

Alignment Scores:

Pred. No.:	0.544	Length:	1773
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	16	Gaps:	0

US-09-823-649A-7 (1-11) x AAR04804 (1-1773)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArglySGlu 11
 |||||
 DB 1303 CTGGCGCAAACTTGAACATTCACGCAAGAA 1335

RESULT 12

AAQ31651
 ID AAQ31651 standard; DNA: 1779 BP.

AC AAQ31651;

DT 07-APR-1993 (first entry)

DE Bacillus caldotenax truncated DNA PolI gene in plasmid pU1205.

KM DNA polymerase; polymerase chain reaction; strain YT-6; DSM406;

KW 5' to 3' exonuclease activity; ds.

OS Synthetic.

FT Key Location/Qualifiers

FT mat_peptide 1..1779

FT /*tag- a /product= N-terminally-deleted.PolI

FT /note= "5' to 3' exonuclease activity deleted"

PN EPP517418-A.

PD 09-DEC-1992.

PF 27-MAY-1992; 92EP-0304763.

PR 03-JUN-1991; 91JP-0157368.

PR 07-NOV-1991; 91JP-0318685.

PR 24-FEB-1992; 92JP-0072090.

PR 25-FEB-1992; 92JP-0073161.

PR 06-APR-1992; 92JP-0112400.

XX (TAKI) TAKARA SHUZO CO LTD.

PI Fujita K, Ishino Y, Kato I, Uemori T;

DR WPI: 1992-40872/50.

DR P-PSDB; AAR28349.

PT Gene for PolI type DNA polymerase and cloning method - for the

PT production of DNA polymerase in high yield.

PS Claim 2; Page 23-24; 30pp; English.

CC A gene coding for DNA polymerase I was isolated from B. caldotenax
 CC YT-6 genomic DNA (see AAQ31650). A series of N-terminal deletions was
 CC prepared to eliminate the region responsible for 5' to 3'
 CC exonuclease activity. Plasmid pU1205 was one of the deletion
 CC plasmids which still expressed PolI activity but for which no 5' to
 CC 3' exonuclease activity was detected. The insert contained in
 CC pU1205 is identical to nucleotides 1090 to 2868 of AAQ31650. E.coli
 CC HB101 transformed with pU1205 produced a heat-resistant DNA
 CC polymerase of mol. wt. 67,000 by SDS-PAGE.
 CC See also AAR28340-R28347 and AAQ31648-Q31649.

CC Sequence 1779 BP; 446 A; 425 C; 529 G; 379 T; 0 other;

XX Alignment Scores:

Pred. No.:

0.547	Length:	1779
Score:	48.00	Matches:
Percent Similarity:	90.91%	Conservative:
Best Local Similarity:	90.91%	Mismatches:
Query Match:	97.96%	Indels:
DB:	13	Gaps:

US-09-823-649A-7 (1-11) x AAQ31651 (1-1779)

QY 1 LeuAlaGlnAsnLeuAsnIle**ArglySGlu 11
 |||||
 DB 1312 TTGGCGCAAACTTGAATATTCACGCAAGAG 1344

RESULT 13

AAQ31651
 ID AAD18444 standard; DNA: 1779 BP.

AC AAD18444;

DT 18-DEC-2001 (first entry)

DE Thermocactinomyces vulgaris DNA polymerase M285 truncated mutant DNA.

KM Reverse transcription; DNA polymerase; Thermocactinomyces vulgaris; TVu;

KW enzyme; amplification; hybridisation; mutant; mutagen; ds.

OS Thermocactinomyces vulgaris.

FT Synthetic.

FT Key Location/Qualifiers

FT cDS 1..1779

FT /*tag- a /product= "Thermocactinomyces vulgaris DNA polymerase

FT M285 truncated mutant"

PN WO200164954-A1.

PD 07-SEP-2001.

PF 28-FEB-2001; 2001WO-US06571.

PR 02-MAR-2000; 2000US-0517871.

PR (PROM-) PROMEGA CORP.

PI Gu T, Huang F, Hartnett JR;

DR WPI: 2001-589872/66.

DR P-PSDB; AAE11057.

PT Reverse transcription of RNA using TVu and Bat DNA polymerase,

PT comprises combining the polymerase, RNA containing sample, a primer and

PT reaction buffer having magnesium ions and reacting the mixture to

PT obtain cDNA -

PS Example 11; Fig 3; 118pp; English.

CC The invention relates to reverse transcription of RNA templates using
 CC thermostable DNA polymerase such as Thermocactinomyces vulgaris (TVu) and
 CC Bacillus stearothermophilus (Bst) DNA polymerase. TVu and Bst DNA
 CC polymerase are useful for reverse transcription reactions which allows
 CC repetitive heating/cooling cycles without the requirement of fresh enzyme
 CC at each cooling step. The invention contemplates single-reaction RT-PCR
 CC wherein reverse transcription and amplification are performed in a single
 CC continuous procedure. Primers, template, nucleoside triphosphates,
 CC appropriate reaction buffer having magnesium ions, reaction conditions,
 CC and polymerase are used in the PCR process, which involves denaturation
 CC of target DNA, hybridisation of primers and synthesis of complementary
 CC strands. The present sequence is Thermocactinomyces vulgaris (TVu) DNA
 CC polymerase M285 truncated mutant DNA. This TVu M285 mutant DNA is
 CC obtained by deleting 852 bases from the 5' end of wild-type TVu DNA
 CC polymerase gene.

SQ Sequence 1779 BP; 430 A; 462 C; 542 G; 345 T; 0 other;

Alignment Scores:

Pred. No.:	0.547	Length:	1779
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-7 (1-11) x AAD18444 (1-1779)

QY 1 leuAlaGlnasnleuAsnle***ArgLysGlu 11
 |||||
 Db 1309 CTGGCGCAAACTTGAACTATCGCGCAAGAA 1341

RESULT 14

AAH47392
 ID AAH47392 standard; DNA; 1779 BP.
 XX
 AC AAH47392;
 XX
 DT 30-NOV-2001 (first entry)
 XX
 DE Tvu DNA polymerase deletion mutant M285 encoding DNA.
 XX
 KW Thermostable; DNA polymerase; hyperthermophilic; DNA synthesis;
 KW Tvu DNA polymerase; nucleic acid amplification; reverse transcription;
 KW deletion mutant; M285; ds.
 XX
 OS Thermophilomycetes vulgaris.
 XX
 FH Key Location/Qualifiers
 FT 1..1779
 FT CDS /tag= a
 FT /product= "Tvu DNA polymerase M285"
 XX
 PN WO200164838-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06436.
 XX
 PR 02-MAR-2000; 2000US-0517439.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Gu T, Huang F, Hartnett JR;
 XX
 DR WPI: 2001-557768/62.
 DR P-PSDB; AAB85950.
 XX
 PT Novel thermostable DNA polymerase derived from Thermophilomycetes
 PT vulgaris for us in many recombinant DNA techniques, including
 PT amplification, reverse transcription and sequencing reactions -
 XX
 PS Claim 17; Fig 5; 113pp; English.
 XS
 CC The invention provides compositions comprising thermostable DNA
 CC polymerases derived from hyperthermophilic eubacteria, in particular,
 CC a purified and isolated Thermophilomycetes vulgaris (Tvu) DNA polymerase
 CC having DNA synthesis activity. It is useful for determining the
 CC nucleotide base sequence of a DNA molecule. Tvu DNA polymerase can be
 CC used for processes of high temperature nucleic acid amplification and
 CC sequencing without substantial loss of DNA synthesis activity. The enzyme
 CC can be used to perform high temperature reverse transcription in the
 CC absence of manganese ions and in high throughput robotically-manipulated
 CC procedures because greater enzymatic stability is retained at room
 CC temperature. Using the Tvu DNA polymerase, reverse transcription
 CC reactions can be conducted at increased temperatures so that secondary
 CC structure is removed or limited. The present sequence represents the
 CC nucleotide sequence encoding a 5'-3' exonuclease deletion mutant form of
 CC Tvu DNA polymerase called M285.

XX
 SQ Sequence 1779 BP; 430 A; 462 C; 542 G; 345 T; 0 other;

Alignment Scores:

Pred. No.:	0.547	Length:	1779
Score:	48.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	97.96%	Indels:	0
DB:	22	Gaps:	0

US-09-823-649A-7 (1-11) x AAH47392 (1-1779)

QY 1 leuAlaGlnasnleuAsnle***ArgLysGlu 11
 |||||
 Db 1309 CTGGCGCAAACTTGAACTATCGCGCAAGAA 1341

RESULT 15

AAH47392
 ID AAH47392 standard; DNA; 2631 BP.
 XX
 AC AAH47392;
 XX
 DT 30-MAY-1996 (first entry)
 XX
 DE B.steartothermophilus DNA polymerase coding sequence.
 XX
 KW Bacillus steartothermophilus; thermostable DNA polymerase; PCR; primer;
 KW amplification; probe; Bacillus caldotenax; exonuclease; recombination;
 KW sequencing; polymerase chain reaction; ss.
 XX
 OS Bacillus steartothermophilus.
 XX
 PN WO9527067-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 30-MAR-1995; 95WO-US04080.
 XX
 PR 24-FEB-1995; 95US-0394232.
 PR 01-APR-1994; 94US-0222612.
 PR 16-SEP-1994; 94US-0307410.
 XX
 PA (GENP-) GEN-PROBE INC.
 XX
 PI Riggs MG, Sivaram M, Tudor SD;
 XX
 DR WPI: 1995-374510/48.
 DR P-PSDB; AAR80137.
 XX
 PT Bacillus steartothermophilus DNA encoding thermostable polymerase -
 PT for nucleic acid sequencing and amplification
 XX
 PS Claim 6; Page 79-82; 135pp; English.
 XS
 CC This is the sequence of the coding region of the gene encoding the
 CC Bacillus steartothermophilus thermostable DNA polymerase. The gene
 CC was isolated using the PCR primers and probes AAT04807-15 which
 CC were based on sequence similarity with the Bacillus caldotenax DNA
 CC polymerase gene. The gene was isolated as two overlapping fragments
 CC of 885 and 1143 bp corresp. to the 3' end and the 3'-5' exonuclease
 CC region of the protein, resp. These amplification fragments were used
 CC to probe Southern gels and isolate the 3' and 5' ends of the gene as
 CC two separate fragments. The complete gene (AAT04801) was subsequently
 CC cloned by recombination of the overlapping fragments. The new DNA
 CC polymerase can be used in primer extension reactions such as DNA
 CC sequencing or polymerase chain reaction.
 XX
 SQ Sequence 2631 BP; 659 A; 676 C; 780 G; 516 T; 0 other;

Alignment Scores:

Pred. No.:	0.869	Length:	2631
Score:	46.00	Matches:	10

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 ; Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec

Title: US-09-823-649a-3
Perfect score: 54
Sequence: 1 LSOELAIPEEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	54	100.0	528 2	US-08-484-956-90 Sequence 90, Appl
2	54	100.0	528 2	US-08-757-653-90 Sequence 90, Appl
3	54	100.0	528 4	US-08-520-946-90 Sequence 90, Appl
4	54	100.0	553 4	US-09-096-399-2 Sequence 2, Appl
5	54	100.0	559 1	US-08-021-623C-6 Sequence 6, Appl
6	54	100.0	559 4	US-09-096-399-4 Sequence 4, Appl
7	54	100.0	560 5	PCT-US95-1418-5 Sequence 5, Appl
8	54	100.0	560 5	PCT-US95-15327-5 Sequence 5, Appl
9	54	100.0	567 5	PCT-US95-1418-4 Sequence 4, Appl
10	54	100.0	597 5	PCT-US95-15327-4 Sequence 4, Appl
11	54	100.0	810 4	US-09-587-856-2 Sequence 2, Appl
12	54	100.0	810 4	US-09-777-537-2 Sequence 2, Appl
13	54	100.0	810 4	US-09-777-538-2 Sequence 2, Appl
14	54	100.0	832 1	US-07-977-434-2 Sequence 2, Appl
15	54	100.0	832 1	US-08-156-020-2 Sequence 2, Appl
16	54	100.0	832 1	US-08-156-020-4 Sequence 2, Appl
17	54	100.0	832 1	US-08-156-020-6 Sequence 6, Appl
18	54	100.0	832 1	US-08-156-020-8 Sequence 8, Appl
19	54	100.0	832 1	US-08-156-020-10 Sequence 10, Appl
20	54	100.0	832 1	US-08-073-384C-4 Sequence 4, Appl
21	54	100.0	832 1	US-08-254-359A-4 Sequence 4, Appl
22	54	100.0	832 1	US-08-483-043-4 Sequence 4, Appl
23	54	100.0	832 1	US-08-483-819-2 Sequence 2, Appl
24	54	100.0	832 2	US-08-481-238-4 Sequence 4, Appl
25	54	100.0	832 2	US-08-471-066B-4 Sequence 4, Appl
26	54	100.0	832 2	US-08-484-956-4 Sequence 4, Appl
27	54	100.0	832 2	US-08-757-653-4 Sequence 4, Appl

28	54	100.0	832 2	US-08-599-491-4 Sequence 4, Appl
29	54	100.0	832 2	US-08-756-386-4 Sequence 4, Appl
30	54	100.0	832 2	US-08-823-516-4 Sequence 4, Appl
31	54	100.0	832 3	US-08-682-853A-4 Sequence 4, Appl
32	54	100.0	832 3	US-08-753-038-4 Sequence 4, Appl
33	54	100.0	832 3	US-08-758-314-4 Sequence 4, Appl
34	54	100.0	832 4	US-09-350-309-4 Sequence 4, Appl
35	54	100.0	832 4	US-08-520-946-4 Sequence 4, Appl
36	54	100.0	832 4	US-08-978-806-2 Sequence 2, Appl
37	54	100.0	832 5	PCT-US91-07035-2 Sequence 2, Appl
38	54	100.0	833 1	US-08-073-384C-8 Sequence 8, Appl
39	54	100.0	833 1	US-08-254-359A-8 Sequence 8, Appl
40	54	100.0	833 1	US-08-483-043-8 Sequence 8, Appl
41	54	100.0	833 1	US-08-481-238-8 Sequence 8, Appl
42	54	100.0	833 2	US-08-471-066B-8 Sequence 8, Appl
43	54	100.0	833 2	US-08-484-956-8 Sequence 8, Appl
44	54	100.0	833 2	US-08-484-956-8 Sequence 8, Appl
45	54	100.0	833 2	US-08-757-653-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-484-956-90
Sequence 90, Application US/08484956
Patent No. 5843654
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROW, MARY ANN D.
APPLICANT: OLDENBURG, MARY C.
APPLICANT: HEISLER, LAURA
TITLE OF INVENTION: DETECTION OF p53 MUTATIONS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAYESSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL J. PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FOS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

;; INFORMATION FOR SEQ ID NO: 90:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 528 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-484-956-90

Query Match 100.0%; Score 54; DB 2; Length 528;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
|||||
Db 374 LSQELAIPIYEE 384

RESULT 2
US-08-757-653-90
; Sequence 7, Application US/08757653
; Patent No. 5843669
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; TITLE OF INVENTION: Thermostable PEN-1 Endonucleases
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-757-653-90

Query Match 100.0%; Score 54; DB 2; Length 528;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
|||||
Db 374 LSQELAIPIYEE 384

RESULT 3
US-08-520-946-90
; Sequence 90, Application US/08520946

;; Patent No. 6372424
;; GENERAL INFORMATION:
;; APPLICANT: BROW, MARY ANN D.
;; APPLICANT: LYAMICHEV, VICTOR I.
;; APPLICANT: OLIVE, DAVID M.
;; TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
;; TITLE OF INVENTION: PATHOGENS
;; NUMBER OF SEQUENCES: 160
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDIEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/520,946
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 528 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-520-946-90

Query Match 100.0%; Score 54; DB 4; Length 528;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
|||||
Db 374 LSQELAIPIYEE 384

RESULT 4
US-09-096-399-2
; Sequence 2, Application US/09096399A
; Patent No. 6130045
; GENERAL INFORMATION:
; APPLICANT: Wurstl, Helmut
; APPLICANT: Qul, Zhi-Hao
; TITLE OF INVENTION: Thermostable Polymerase
; FILE REFERENCE: CLON-007
; CURRENT APPLICATION NUMBER: US/09/096,399A
; CURRENT FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinantly engineered mutant
US-09-096-399-2

Query Match 100.0%; Score 54; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
|||||
Db 399 LSOELAIPIYEE 409

RESULT 5

US-08-021-623C-6
Sequence 6, Application US/08021623C
Patent No. 5436149
GENERAL INFORMATION:
APPLICANT: Barnes Ph.D., Wayne M.
TITLE OF INVENTION: Thermostable DNA polymerase with
TITLE OF INVENTION: enhanced thermostability and enhanced length and
TITLE OF INVENTION: efficiency of primer extension
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Senninger, Powers, Leavitt and Roedel
STREET: One Metropolitan Square, 16th floor
City: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 5.25", 360 kb.
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln 1.0, v1.25; EDIX; Wordperfect.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,623C
FILING DATE: 19-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Blosser, G. Hatley
REGISTRATION NUMBER: 33,650
REFERENCE/DOCKET NUMBER: WNB4900
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314/231-5400
TELEFAX: 314/231-4342
TELEX: 6502697583 KCI
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-623C-6

Query Match 100.0%; Score 54; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
|||||
Db 400 LSOELAIPIYEE 410

RESULT 6

US-09-096-399-4
Sequence 4, Application US/09096399A
Patent No. 6130045
GENERAL INFORMATION:
APPLICANT: Wurst, Helmut
APPLICANT: Qui, Zhi-Hao
TITLE OF INVENTION: Thermostable Polymerase
FILE REFERENCE: CLON-007
CURRENT APPLICATION NUMBER: US/09/096,399A
TITLE OF INVENTION: Thermus Flavus DNA Polymerase
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 559
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Recombinantly engineered mutant
US-09-096-399-4

Query Match 100.0%; Score 54; DB 4; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
|||||
Db 405 LSOELAIPIYEE 415

RESULT 7

PCT-US95-14418-5
Sequence 5, Application PC/TUS9514418
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA Encoding a Thermostable DNA Polymerase Enzyme
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
City: Chicago
STATE: Illinois
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14418
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 28003/32330
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-14418-5

Query Match 100.0%; Score 54; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
|||||
Db 406 LSOELAIPIYEE 416

RESULT 8

PCT-US95-15327-5
Sequence 5, Application PC/TUS9515327
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Biologically Active Fragments of
TITLE OF INVENTION: Thermus Flavus DNA Polymerase
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive

```

? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/15327
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Gass, David A.
? REGISTRATION NUMBER: 38,153
? REFERENCE/DOCKET NUMBER: 28003/31716
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6300
? TELEFAX: 312/474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 560 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
PCT-US95-15327-5

Query Match
Best Local Similarity 100.0%; Score 54; DB 5; Length 560;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11 --
Db 406 LSQELAIPIYEE 416

RESULT 9
PCT-US95-14418-4
? Sequence 4, Application PC/TUS9514418
? GENERAL INFORMATION:
? APPLICANT:
? TITLE OF INVENTION: DNA Encoding a Thermostable DNA Polymerase Enzyme
? NUMBER OF SEQUENCES: 51
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
? STREET: 6300 Sears Tower, 233 South Wacker Drive
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/14418
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Gass, David A.
? REGISTRATION NUMBER: 38,153
? REFERENCE/DOCKET NUMBER: 28003/31716
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6300
? TELEFAX: 312/474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 597 amino acids
? TYPE: amino acid
? TOPOLOGY: linear

```

```

? MOLECULE TYPE: protein
PCT-US95-14418-4

Query Match
Best Local Similarity 100.0%; Score 54; DB 5; Length 597;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
Db 443 LSQELAIPIYEE 453

RESULT 10
PCT-US95-15327-4
? Sequence 4, Application PC/TUS9515327
? GENERAL INFORMATION:
? APPLICANT:
? TITLE OF INVENTION: Biologically Active Fragments of
? TITLE OF INVENTION: Thermus Flavus DNA Polymerase
? NUMBER OF SEQUENCES: 51
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
? STREET: 6300 Sears Tower, 233 South Wacker Drive
? CITY: Chicago
? STATE: Illinois
? COUNTRY: United States of America
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/15327
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Gass, David A.
? REGISTRATION NUMBER: 38,153
? REFERENCE/DOCKET NUMBER: 28003/31716
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6300
? TELEFAX: 312/474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 597 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
PCT-US95-15327-4

Query Match
Best Local Similarity 100.0%; Score 54; DB 5; Length 597;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
Db 443 LSQELAIPIYEE 453

RESULT 11
US-09-587-856-2
? Sequence 2, Application US/09587856
? Patent No. 6214557
? GENERAL INFORMATION:
? APPLICANT: Washington University
? TITLE OF INVENTION: COLD SENSITIVE MUTANT DNA POLYMERASES
? FILE REFERENCE: WSHO 2009
? CURRENT APPLICATION NUMBER: US/09/587,856
? CURRENT FILING DATE: 2000-06-06
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 2

```


LENGTH: 810
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-587-856-2

Query Match 100.0%; Score 54; DB 4; Length 810;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSOELAIPEE 11
Db 678 LSOELAIPEE 688

RESULT 12
US-09-777-537-2
Sequence 2, Application US/09777537
Patent No. 6316202
GENERAL INFORMATION:
APPLICANT: Barnes, Wayne M
APPLICANT: Kornekchlev, Milko B
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING COLD SENSITIVE MUTANT DNA
FILE REFERENCE: MSHU 2009.1
CURRENT APPLICATION NUMBER: US/09/777,537
CURRENT FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 09/587,856
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 810
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-777-537-2

Query Match 100.0%; Score 54; DB 4; Length 810;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSOELAIPEE 11
Db 678 LSOELAIPEE 688

RESULT 13
US-09-777-538-2
Sequence 2, Application US/09777538
Patent No. 6333159
GENERAL INFORMATION:
APPLICANT: Barnes, Wayne M
APPLICANT: Kornekchlev, Milko B
TITLE OF INVENTION: COLD SENSITIVE MUTANT DNA POLYMERASIS AND METHODS OF
FILE REFERENCE: MSHU 2009.2
CURRENT APPLICATION NUMBER: US/09/777,538
CURRENT FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 09/587,856
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 810
TYPE: PRT
ORGANISM: Thermus aquaticus
US-09-777-538-2

Query Match 100.0%; Score 54; DB 4; Length 810;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LSOELAIPEE 11
Db 678 LSOELAIPEE 688

Db 678 LSOELAIPEE 688

RESULT 14
US-07-977-434-2
Sequence 2, Application US/07977434
Patent No. 546591
GENERAL INFORMATION:
APPLICANT: Goltand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSER: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: WordPerfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseri
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 546591 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-977-434-2

Query Match 100.0%; Score 54; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
|||||
DB 678 LSQELAIPEE 688

RESULT 15

US-08-156-020-2
Sequence 2, Application US/08156020
Patent No. 5474920
GENERAL INFORMATION:
APPLICANT: Moses M.D., Robb E.
TITLE OF INVENTION: Modified Thermo-Resistant DNA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegettl & Wilcoff
STREET: 10 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/156,020
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Greenfield Ph.D., Michael S.
REGISTRATION NUMBER: 37,142
REFERENCE/DCKET NUMBER: 93,413
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)715-1000
TELEFAX: (312)715-1234
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-156-020-2

Query Match 100.0%; Score 54; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
|||||
DB 678 LSQELAIPEE 688

Search completed: January 15, 2003, 12:38:22
Job time : 9.28571 secs


```
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP41006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus aquaticus
US-09-823-649a-8
```

```
Query Match
Best Local Similarity 100.0%; Score 54; DB 10; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSOELAIPEEE 11
Db 1 LSOELAIPEEE 11
```

```
RESULT 3
US-09-823-649a-10
; Sequence 10, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP41006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus thermophilus
US-09-823-649a-10
```

```
Query Match
Best Local Similarity 100.0%; Score 54; DB 10; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSOELAIPEEE 11
Db 1 LSOELAIPEEE 11
```

```
RESULT 4
US-09-823-649a-11
; Sequence 11, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
```

```
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP41006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus sp. 205
US-09-823-649a-11
```

```
Query Match
Best Local Similarity 100.0%; Score 54; DB 10; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSOELAIPEEE 11
Db 1 LSOELAIPEEE 11
```

```
RESULT 5
US-09-823-649a-13
; Sequence 13, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
; FILE REFERENCE: RP41006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus caldophilus
US-09-823-649a-13
```

```
Query Match
Best Local Similarity 100.0%; Score 54; DB 10; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSOELAIPEEE 11
Db 1 LSOELAIPEEE 11
```

```
RESULT 6
US-10-033-297-4
; Sequence 4, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; Sequential Invasive Cleavages
```

```

NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020187486A1 Relevant
TOPOLOGY: No. US20020187486A1 Relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-033-297-4

Query Match
Best Local Similarity 100.0%; Score 54; LB 9; Length 832;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIYEE 11
|||||
Db 678 LSQELAIYEE 688

RESULT 7
US-09-972-834-2
Sequence 2, Application US/09972834
Publication No. US20020192663A1
GENERAL INFORMATION:
APPLICANT: Loeb, Lawrence A.
Hood, Leroy
Suzuki, Motohshi
TITLE OF INVENTION: Thermostable Polymerases Having Altered
Fidelity and Methods of Identifying and Using Same
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL & FLORES LLP

```

```

STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92121

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,834
FILING DATE: 04-OCT-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/978,806
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UM 2873
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)535-9001
TELEFAX: (619)535-8949
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-972-834-2

Query Match
Best Local Similarity 100.0%; Score 54; DB 9; Length 832;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIYEE 11
|||||
Db 678 LSQELAIYEE 688

RESULT 8
US-10-071-505-1
Sequence 1, Application US/10071505
Publication No. US20020197211A1
GENERAL INFORMATION:
APPLICANT: Davis, Maria
APPLICANT: Nelson, John
APPLICANT: Kumar, Shiv
APPLICANT: Finn, Patrick J.
APPLICANT: Nampalli, Salyam
APPLICANT: Flick, Parke
TITLE OF INVENTION: TAQ DNA Polymerase Having an Amino Acid Substitution at
File Reference: PB9944
FILE REFERENCE: PB9944
CURRENT APPLICATION NUMBER: US/10/071,505
PRIOR APPLICATION NUMBER: 2002-02-08
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 1
LENGTH: 832
TYPE: PRT
ORGANISM: Thermus aquaticus
US-10-071-505-1

Query Match
Best Local Similarity 100.0%; Score 54; DB 9; Length 832;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LSOELAIPEE 11
|||||
Db 678 LSOELAIPEE 688

RESULT 9
US-10-081-806-4
; Sequence 4, Application US/10081806
; Publication No. US20020197623A1
; GENERAL INFORMATION:
; APPLICANT: Prudent, James R.
; Hall, Jeff G.
; Lyamichev, Victor I.
; TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/081,806
; FILING DATE: 22-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,386
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02564
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-081-806-4

Query Match 100.0%; Score 54; DB 9; Length 832;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPEE 11
|||||
Db 678 LSOELAIPEE 688

RESULT 10
US-10-033-297-8
; Sequence 8, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; Lyamichev, Victor I.
; Mast, Andrea L.
; Brow, Mary Ann D.

; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/033,297
; FILING DATE: 12-No. US20020187486A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/350,597
; FILING DATE: 09-JUL-1999
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 833 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-033-297-8

Query Match 100.0%; Score 54; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPEE 11
|||||
Db 678 LSOELAIPEE 688

RESULT 11
US-10-033-297-66
; Sequence 66, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; Lyamichev, Victor I.
; Mast, Andrea L.
; Brow, Mary Ann D.

TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequential Invasive Cleavages

NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 66:
US-10-033-297-66
Query Match 100.0%; Score 54; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSEQALPYEE 11
Db 679 LSEQALPYEE 689
RESULT 12
US-10-033-297-69
Sequence 69, Application US/1003297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 69:
US-10-033-297-69
Query Match 100.0%; Score 54; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSEQALPYEE 11
Db 679 LSEQALPYEE 689
RESULT 13
US-10-033-297-71
Sequence 71, Application US/1003297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyamichev, Victor I.
Mast, Andrea L.
Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSER: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California

COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <unknown>
PRIORITY INFORMATION:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 71:
US-10-033-297-71
Query Match 100.0%; Score 54; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSQELAIPEE 11
DB 679 LSQELAIPEE 689
RESULT 14
US-10-081-806-8
Sequence 8, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyanichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <unknown>
PRIORITY INFORMATION:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-081-806-8
Query Match 100.0%; Score 54; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSQELAIPEE 11
DB 678 LSQELAIPEE 688
RESULT 15
US-10-033-297-6
Sequence 6, Application US/10033297
Publication No. US20020187486A1
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
Lyanichev, Victor I.
Brow, Mary Ann D.
West, Andrea D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
Sequences
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/033,297
FILING DATE: 12-NOV-1996
CLASSIFICATION: <unknown>
PRIORITY INFORMATION:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072


```
?
?
? FILING DATE: 21-JAN-1997
? APPLICATION NUMBER: US 08/759,038
? FILING DATE: 02-DEC-1996
? APPLICATION NUMBER: US 08/758,314
? FILING DATE: 02-DEC-1996
? APPLICATION NUMBER: US 08/756,386
? FILING DATE: 29-NOV-1996
? APPLICATION NUMBER: US 08/682,853
? FILING DATE: 12-JUL-1996
? APPLICATION NUMBER: US 08/599,491
? FILING DATE: 24-JAN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Ingolia, Diane E.
? REGISTRATION NUMBER: 40,027
? REFERENCE/DOCKET NUMBER: FORS-02736
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 6:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 834 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-033-297-6

Query Match 100.0%; Score 54; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. NO. 0.02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSGELAIPIYEE 11
|||||
Db 680 LSGELAIPIYEE 690
```

Search completed: January 15, 2003, 12:39:50
Job time : 10.8571 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 12:36:21; Search time 26 Seconds
(without alignments)
129.748 Million cell updates/sec

Title: US-09-823-649a-2
Perfect score: 47
Sequence: 1 LSXELXIPYEE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 862724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODE=frame_plus_p2n.model -DEV=x1h
-O=/cgn2_1/USPTO.spool/US09823649/unat.14012003.151002.29127/app.query.fasta_1.1393
-DB=Issued_Patents_NA -QEXT=fastap -SUFFIX=rni -MINMATCH=0.1 -IOCPCL=0
-LOOPEXT=0 -UNITITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIG=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTEXT=PCT -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649 -QCGN_1.1.35 -Grunat.14012003.151002.29127 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database: Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PT08.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	93.6	1600	US-08-073-384C-12	Sequence 12, Appl
2	44	93.6	1600	US-08-254-359A-12	Sequence 12, Appl
3	44	93.6	1600	US-08-483-043-12	Sequence 12, Appl
4	44	93.6	1600	US-08-481-238-12	Sequence 12, Appl
5	44	93.6	1600	US-08-471-066B-12	Sequence 12, Appl
6	44	93.6	1600	US-08-484-956-12	Sequence 12, Appl
7	44	93.6	1600	US-08-757-653-12	Sequence 12, Appl
8	44	93.6	1600	US-08-599-491-12	Sequence 12, Appl
9	44	93.6	1600	US-08-756-386-12	Sequence 12, Appl
10	44	93.6	1600	US-08-823-516-12	Sequence 12, Appl
11	44	93.6	1600	US-08-682-853A-12	Sequence 12, Appl
12	44	93.6	1600	US-08-759-038-12	Sequence 12, Appl

13	44	93.6	1600	US-08-758-314-12	Sequence 12, Appl
14	44	93.6	1600	US-09-350-309-12	Sequence 12, Appl
15	44	93.6	1600	US-08-520-946-12	Sequence 12, Appl
16	44	93.6	1682	US-09-026-399-1	Sequence 1, Appl
17	44	93.6	1686	US-08-648-657-1	Sequence 1, Appl
18	44	93.6	1686	US-08-648-657-14	Sequence 14, Appl
19	44	93.6	1686	US-08-648-657-2	Sequence 2, Appl
20	44	93.6	1696	US-09-096-399-3	Sequence 3, Appl
21	44	93.6	1794	PCT-US95-14418-3	Sequence 3, Appl
22	44	93.6	1794	PCT-US95-15327-3	Sequence 3, Appl
23	44	93.6	2493	US-07-977-434-5	Sequence 5, Appl
24	44	93.6	2493	US-08-458-819-5	Sequence 5, Appl
25	44	93.6	2493	PCT-US91-07035-5	Sequence 5, Appl
26	44	93.6	2496	US-08-073-384C-2	Sequence 2, Appl
27	44	93.6	2496	US-08-471-066B-2	Sequence 2, Appl
28	44	93.6	2496	US-08-484-956-2	Sequence 2, Appl
29	44	93.6	2496	US-08-757-653-2	Sequence 2, Appl
30	44	93.6	2496	US-08-599-491-2	Sequence 2, Appl
31	44	93.6	2496	US-08-823-516-2	Sequence 2, Appl
32	44	93.6	2496	US-08-823-516-2	Sequence 2, Appl
33	44	93.6	2496	US-08-823-516-2	Sequence 2, Appl
34	44	93.6	2496	US-08-823-516-2	Sequence 2, Appl
35	44	93.6	2496	US-08-823-516-2	Sequence 2, Appl
36	44	93.6	2496	US-08-823-516-2	Sequence 2, Appl
37	44	93.6	2496	US-08-759-038-2	Sequence 2, Appl
38	44	93.6	2496	US-08-758-314-2	Sequence 2, Appl
39	44	93.6	2496	US-09-350-309-2	Sequence 2, Appl
40	44	93.6	2496	US-08-520-946-2	Sequence 2, Appl
41	44	93.6	2499	US-07-977-434-1	Sequence 1, Appl
42	44	93.6	2499	US-08-458-819-1	Sequence 1, Appl
43	44	93.6	2499	US-07-602-848E-1	Sequence 1, Appl
44	44	93.6	2499	US-09-587-856-1	Sequence 1, Appl
45	44	93.6	2499	US-09-777-537-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-073-384C-12
Sequence 12, Application US/08073384C
Patent No. 5541311
GENERAL INFORMATION:
APPLICANT: Dahlberg, James E.
APPLICANT: Lymanchev, Victor I.
APPLICANT: Brown, Mary Ann D.
TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVESTOCK, MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073.384C
FILING DATE: 04-JUN-1993
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-00613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410

TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-073-384C-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 1 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-073-384C-12 (1-1600)

QY 1 Leuser***Gluleu***1leProTyrgluGlu 11
||||| ||||| ||||| ||||| |||||
Db 1133 CTCCTCCAGAGACTACCATCCCTTACGAGAG 1165

RESULT 2
US-08-254-359A-12

; Sequence 12, Application US/08254359A
; Patent No. 5614402

; GENERAL INFORMATION:

; APPLICANT: DAHLBERG, JAMES E.

; APPLICANT: LYAMICHEV, VICTOR I.

; APPLICANT: BROW, MARY ANN D.

; TITLE OF INVENTION: 5' NUCLEASIS DERIVED FROM THERMOSTABLE

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HAVERSOCK, MEDLEN & CARROLL

; STREET: 220 MONTGOMERY STREET, SUITE 2200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/254,359A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/073,384

; FILING DATE: 06-JUN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/986,330

; FILING DATE: 07-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: CARROLL, PETER G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: F08S-01000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-254-359A-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 1 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-254-359A-12 (1-1600)

QY 1 Leuser***Gluleu***1leProTyrgluGlu 11
||||| ||||| ||||| ||||| |||||
Db 1133 CTCCTCCAGAGACTACCATCCCTTACGAGAG 1165

RESULT 3

US-08-483-043-12

; Sequence 12, Application US/08483043
; Patent No. 5691142

; GENERAL INFORMATION:

; APPLICANT: Dahlberg, James E.

; APPLICANT: Lyamichev, Victor I.

; APPLICANT: Brow, Mary Ann D.

; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HAVERSOCK, MEDLEN & CARROLL

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,043

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/073,384

; FILING DATE: 04-JUN-1993

; APPLICATION NUMBER: US 07/986,330

; FILING DATE: 07-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Carroll, Peter G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: F08S 00613

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/705-8410

; TELEFAX: 415/397-8338

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-483-043-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 1 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-483-043-12 (1-1600)

QY 1 Leuser***Gluleu***1leProTyrgluGlu 11

```
Db      1133 CTCCTCCAGAGAGCTACGACATCCCTTACGAGAG 1165
RESULT 4
US-08-481-238-12
; Sequence 12, Application US/08481238
; Patent No. 5795763
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: BROW, MARY ANN D.
; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,238
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: PORS-01798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-481-238-12

Alignment Scores:
Pred. No.: 0.596      Length: 1600
Score: 44.00      Matches: 9
Percent Similarity: 81.82%      Conservative: 0
Best Local Similarity: 81.82%      Mismatches: 2
Query Match: 93.62%      Indels: 0
DB: 1      Gaps: 0

US-09-823-649A-2 (1-11) x US-08-481-238-12 (1-1600)
QY      1 Leuser**GluLeu**lleProTyrGlnGlu 11
Db      1133 CTCCTCCAGAGAGCTACGACATCCCTTACGAGAG 1165

RESULT 5
US-08-471-066B-12
; Sequence 12, Application US/08471066B
; Patent No. 5837450
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: BROW, MARY ANN D.
; TITLE OF INVENTION: 5' Nucleases Derived From Thermostable
; Nucleases
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,066B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PORS-01800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-471-066B-12

Alignment Scores:
Pred. No.: 0.596      Length: 1600
Score: 44.00      Matches: 9
Percent Similarity: 81.82%      Conservative: 0
Best Local Similarity: 81.82%      Mismatches: 2
Query Match: 93.62%      Indels: 0
DB: 2      Gaps: 0

US-09-823-649A-2 (1-11) x US-08-471-066B-12 (1-1600)
QY      1 Leuser**GluLeu**lleProTyrGlnGlu 11
Db      1133 CTCCTCCAGAGAGCTACGACATCCCTTACGAGAG 1165

RESULT 6
US-08-484-956-12
; Sequence 12, Application US/08484956
; Patent No. 5843654
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: BROW, MARY ANN D.
; APPLICANT: OLDENBURG, MARY C.
; APPLICANT: HEISTER, LAURA
; TITLE OF INVENTION: DETECTION OF P53 MUTATIONS
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAYESTOCK, MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL J, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 397-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-956-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-484-956-12 (1-1600)
QY 1 Leuser***Gluleu***IleProTyGluGlu 11
Db 1133 CTCCTCCAGGAGCTAGCCATCCCTTAGCAGCAG 1165

RESULT 7
US-08-757-653-12
Sequence 12, Application US/08/57653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-653-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-757-653-12 (1-1600)
QY 1 Leuser***Gluleu***IleProTyGluGlu 11
Db 1133 CTCCTCCAGGAGCTAGCCATCCCTTAGCAGCAG 1165

RESULT 8
US-08-599-491-12
Sequence 12, Application US/08/599491
Patent No. 5846717
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: GROTELOESCHEN HALL, JEFF S.
APPLICANT: LYAMICHEV, VICTOR
APPLICANT: OLIVE, DAVID M.
APPLICANT: PROBERT, JAMES R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSER: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,491
FILING DATE: 23-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: P-40,027
REFERENCE/DOCKET NUMBER: FORS-01802
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1600 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-599-491-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-599-491-12 (1-1600)

Qy 1 Leuser***Glueu***lleProtyrGluGlu 11
||||| ||||| ||||| ||||| |||||
Db 1133 CTCGCCAGAGGAGCTAGCATCCCTTACGAGAG 1165

RESULT 9
US-08-756-386-12
; Sequence 12, Application US/08756386
; Patent No. 5985557
; GENERAL INFORMATION:
; APPLICANT: Prudent, James R.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
; NUMBER OF SEQUENCES: .65*
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,386
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02564
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-756-386-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600

Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-756-386-12 (1-1600)

Qy 1 Leuser***Glueu***lleProtyrGluGlu 11
||||| ||||| ||||| ||||| |||||
Db 1133 CTCGCCAGAGGAGCTAGCATCCCTTACGAGAG 1165

RESULT 10
US-08-823-516-12
; Sequence 12, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea I.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 28-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,353
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-823-516-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-2 (1-11) x US-08-823-516-12 (1-1600)

QY 1 Leuser***Gluleu***IleProTyrGluGln 11

DB 1133 CTCCTCCAGAGCTAGCCTCCCTTACGAGAG 1165

RESULT 11

US-08-682-853a-12
Sequence 12, Application US/08682853A
Patent No. 6001567

GENERAL INFORMATION:

APPLICANT: Brow, Mary Ann D.
APPLICANT: Grotelueschen Hall, Jeff S.

APPLICANT: Lymichev, Victor

APPLICANT: Olive, David M.

APPLICANT: Prudent, James R.

TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESS: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/682,853A

FILING DATE: 12-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/599,491

FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FOS-02306

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1600 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-682-853a-12

Alignment Scores:

Pred. No.: 0.596

Score: 44.00

Percent Similarity: 81.82%

Best Local Similarity: 81.82%

Query Match: 93.62%

DB: 3

US-09-823-649a-2 (1-11) x US-08-682-853a-12 (1-1600)

QY 1 Leuser***Gluleu***IleProTyrGluGln 11

DB 1133 CTCCTCCAGAGCTAGCCTCCCTTACGAGAG 1165

RESULT 12

US-08-759-038-12

Sequence 12, Application US/08759038

Patent No. 6090543

GENERAL INFORMATION:

APPLICANT: Prudent, James R.

APPLICANT: Hall, Jeff G.

APPLICANT: Lymichev, Victor I.

APPLICANT: Brow, Mary Ann D.

APPLICANT: Dahlberg, James R.

TITLE OF INVENTION: Cleavage of Nucleic Acids

NUMBER OF SEQUENCES: 134

CORRESPONDENCE ADDRESS:

ADDRESS: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/759,038

FILING DATE: 02-DEC-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/

FILING DATE: 29-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/682,853

FILING DATE: 12-JUL-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/599,491

FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FOS-02574

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1600 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-759-038-12

Alignment Scores:

Pred. No.: 0.596

Score: 44.00

Percent Similarity: 81.82%

Best Local Similarity: 81.82%

Query Match: 93.62%

DB: 3

US-09-823-649a-2 (1-11) x US-08-759-038-12 (1-1600)

QY 1 Leuser***Gluleu***IleProTyrGluGln 11

DB 1133 CTCCTCCAGAGCTAGCCTCCCTTACGAGAG 1165

RESULT 13

US-08-758-314-12

```
Sequence 12, Application US/08758314
Patent No. 6090606
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Improved Cleavage Agents
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,314
FILING DATE: 02-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02575
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-758-314-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0
DB: 3

US-09-823-649a-2 (1-11) x US-08-758-314-12 (1-1600)
QY 1 leuser***GluLeu***lleProTyrGluGlu 11
Db 1133 CTCCTCCAGGAGCTAGCCATCTCCCTACGAGAG 1165

RESULT 14
US-09-350-309-12
Sequence 12, Application US/09350309
Patent No. 6348314
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ. ID NO: 12:
US-09-350-309-12

Alignment Scores:
Pred. No.: 0.596 Length: 1600
Score: 44.00 Matches: 9
Percent Similarity: 81.82% Conservative: 0
Best Local Similarity: 81.82% Mismatches: 2
Query Match: 93.62% Indels: 0
Gaps: 0
DB: 4

US-09-823-649a-2 (1-11) x US-09-350-309-12 (1-1600)
QY 1 leuser***GluLeu***lleProTyrGluGlu 11
Db 1133 CTCCTCCAGGAGCTAGCCATCTCCCTACGAGAG 1165

RESULT 15
US-08-520-946-12
Sequence 12, Application US/08520946
Patent No. 6372424
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: OLIVE, DAVID M.
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```



```

;
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/520,946
;   FILING DATE:
;   CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  CARROLL, PETER G.
;   REGISTRATION NUMBER:  32,837
;   REFERENCE/DOCKET NUMBER:  FORS-01756
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (415) 705-8410
;   TELEFAX:  (415) 397-8338
;   INFORMATION FOR SEQ ID NO: 12:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1600 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   US-08-520-946-12

```

```

Alignment Scores:
Pred. No.:      0.596      Length:      1600
Score:          44.00      Matches:      9
Percent Similarity: 81.82%      Conservative: 0
Best Local Similarity: 81.82%      Mismatches: 2
Query Match:      93.62%      Indels:    0
DB:               4      Gaps:          0

```

US-09-823-649a-2 (1-11) x US-08-520-946-12 (1-1600)

```
Oy      1  leuser**Gluleu**I*eproTYrGlucIu 11
          |||||  |||||  |||||  |||||  |||||
DB      1133 CTCCTCCAGAGACTAGCCATCCCTTACGAGAG 1165

```

Search completed: January 15, 2003, 12:58:39
Job time : 29 secs

PT using a mutant thermoactive DNA polymerase -
XX
XX Disclosure; Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate deoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provides a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

SQ Sequence 11 AA;

Query Match 98.0%; Score 48; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0025;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAONLNXRKE 11
Db 1 LAONLNXRKE 11

RESULT 2

ID AAM48270 standard; Peptide: 11 AA.

AC AAM48270;

DT 25-MAR-2002 (first entry)

DE Native DNA polymerase motif #19.

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KW dye; amplification.

OS Bacillus stearothermophilus.

PN EPI152062-A2.

PD 07-NOV-2001.

PF 12-APR-2001; 2001EP-0109341.

PR 18-APR-2000; 2000US-198336P.

PA (HOFF) HOFFMANN LA ROOHE & CO AG F.

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schenbunnert NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermoactive DNA polymerase -
XX Disclosure; Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA

CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate deoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provides a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

SQ Sequence 11 AA;

Query Match 98.0%; Score 48; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0025;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LAONLNXRKE 11
Db 1 LAONLNXRKE 11

RESULT 3

ID AAR80139 standard; Protein; 587 AA.

AC AAR80139;

DT 30-MAY-1996 (first entry)

DE B.stearothermophilus DNA polymerase variant Bst2.

KW Bacillus stearothermophilus; thermostable DNA polymerase; PCR; primer;
KW amplification; probe; Bacillus caldoleax; exonuclease; recombination;
KW sequencing; polymerase chain reaction.

OS Synthetic.

PN W09527067-A1.

PD 12-OCT-1995.

PF 30-MAR-1995; 95WO-US04080.

PR 24-FEB-1995; 95US-0394232.

PR 01-APR-1994; 94US-0222612.

PR 16-SEP-1994; 94US-0307410.

PA (GENP-) GEN-PROBE INC.

PI Riggs MG, Sivaram M, Tudor SD;

PI WPI; 1995-373510/48.

DR N-PSDB; AAT04802.

XX Claim 9; Page 83-85; 135pp; English.

CC This is the amino acid sequence of the modified Bacillus
CC stearothermophilus thermostable DNA polymerase clone Bst2 which lacks
CC the complete 5'-3' exonuclease domain. The corresp. gene was constructed
CC by ligating the HindIII-SalI fragment from the clone pGemBst2.1st contg.
CC 3' end of the gene with the StyI-HindIII fragment of the clone
CC pGemBst5 end contg. the 5' end of the DNA polymerase gene. The
CC StyI-HindIII fragment covers the mid-region of the gene up to and
CC including the sequences encoding the 3'-5' exonuclease but not the 5'-3'
CC exonuclease domain. The remainder of the 5' end of the gene was
CC reconstituted by the ligated and extended oligonucleotides AAT04816-17.
CC The modified DNA polymerase can be used in primer extension reactions
CC such as DNA sequencing or polymerase chain reaction.

SQ Sequence 587 AA;
 Query Match 98.0%; Score 48; DB 16; Length 587;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAQNUNIXRKE 11
 |||||
 Db 432 LAQNUNIXRKE 442

RESULT 4
 AAW30093
 ID AAW30093 standard; Protein; 587 AA.
 AC AAW30093;
 XX
 XX 14-APR-1998 (first entry)
 DE Bacillus stearothermophilus DNA polymerase (peptides 1-587).
 XX
 XX Bacillus stearothermophilus; DNA polymerase; DNA amplification;
 KW DNA replication; exonuclease; proof reading.
 XX
 OS Bacillus stearothermophilus.
 PN JP09220087-A.
 PD 26-AUG-1997.
 PF 17-MAY-1996; 96JP-0160402.
 PR 03-MAY-1996; 960S-0642684.
 PR 18-OCT-1995; 950S-0544643.
 XX
 PA (HONG/) HONG G. F.
 PA (ZHAI/) ZHAI F.
 PA (SHAN-) SHANGHAI INST BIOCHEMISTRY CHINESE ACAD.
 PI Hong GF, Huang W, Zhai F;
 DR MPI: 1997-474304/44.
 DR N-PSDB; AAT90819.
 XX
 PT DNA polymerase from Bacillus stearothermophilus - used in DNA
 PT amplification and sequencing methods having lower frequency of
 PT mismatch
 XX
 PS Claim 1; Page 2; 32pp; Japanese.
 XX
 CC This is the peptide sequence (587 amino acids) of the Bacillus
 CC stearothermophilus strain DNA polymerase. The heat stable DNA polymerase
 CC can be used in a method for the replication of a DNA chain. It can also
 CC be used in a method for the determination of the sequence of DNA chain.
 CC As the DNA polymerase has a proof reading 3'-5' exonuclease activity,
 CC possible mismatches of base pairing that occurs in current methods of DNA
 CC sequence determination of a DNA clone can be avoided.
 XX
 SQ Sequence 587 AA;
 Query Match 98.0%; Score 48; DB 18; Length 587;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAQNUNIXRKE 11
 |||||
 Db 432 LAQNUNIXRKE 442

RESULT 5
 AAW10597
 ID AAW10597 standard; Protein; 587 AA.

XX
 AC AAW10597;
 XX
 DE 07-MAY-1997 (first entry)
 DE Bst DNA polymerase large fragment.
 XX
 XX DNA polymerase; Bst Pol I; thermostable enzyme; exonuclease;
 KW DNA sequencing; strand displacement amplification.
 XX
 OS Bacillus stearothermophilus.
 PN EF57100-A1.
 PD 05-FEB-1997.
 PF 31-JUL-1996; 96EP-0202169.
 PR 02-AUG-1995; 950S-0510215.
 PA (NEWE) NEW ENGLAND BIOLABS INC.
 PI Aliotta JM, Kong H, Pelletier JJ;
 DR MPI: 1997-111048/11.
 DR N-PSDB; T3630.
 PT DNA encoding recombinant thermophilic Bacillus stearothermophilus
 PT DNA polymerase I - substantially free of 3'-5', and opt. also 5'-3'
 PT exonuclease activity
 XX
 PS Disclosure; Page 12-16; 34pp; English.
 CC
 CC Bacillus stearothermophilus DNA polymerase (Bst Pol I) large
 CC fragment (AAW10597) lacks the N-terminal 3'-5' exonuclease domain of
 CC native Bst Pol I. It is encoded by a DNA construct (AAT36360) obtd.
 CC by PCR cloning of genomic DNA using a primer (AAT60826) based on the
 CC large fragment N-terminal sequence (AAW10597). Recombinant Bst Pol I
 CC large fragment can be produced in transformed host cells, esp. as a
 CC fusion protein with majose binding protein. It is thermophilic,
 CC retaining its activity at 65 deg, and is useful for nucleic acid
 CC sequencing and strand displacement amplification.
 XX
 SQ Sequence 587 AA;
 Query Match 98.0%; Score 48; DB 18; Length 587;
 Best Local Similarity 90.9%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LAQNUNIXRKE 11
 |||||
 Db 432 LAQNUNIXRKE 442

RESULT 6
 AAY91931
 ID AAY91931 standard; Protein; 587 AA.
 AC AAY91931;
 XX
 XX 19-JUL-2000 (first entry)
 DE B. stearothermophilus 320 wild type DNA polymerase.
 XX
 XX Modified DNA polymerase; DNA sequencing; 3' to 5' exonuclease activity;
 KW random primer labeling; site-directed mutagenesis; wild type.
 XX
 OS Bacillus stearothermophilus.
 PN
 FT Key location/Qualifiers
 FT Misc-difference 341..343
 FT /note= "modified to Thr-Pro-Leu in AAY91930"
 FT Misc-difference 421

XX 25-MAR-1997. PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
PD XX
XX 06-JUN-1994; 94US-0254359. PI Brow MAD, Dahlberg JE, Lyamichev VI;
PF Olive DM, Prudent JR;
XX 07-DEC-1992; 92US-0986330. DR WPI; 1997-393613/36.
PR 04-JUN-1993; 93US-0073384. XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC. XX
XX PA
PI Brow MAD, Dahlberg JE, Lyamichev VI;
XX WPI; 1997-201481/18. XX
XX
XX Thermostable 5' nuclease derived from thermostable polymerase - has
PT reduced synthetic activity useful in nucleic acid detection assays
XX
XX Example 2A; Columns 79-82; 93pp; English. XX
XX
XX The present sequence, clone 4F, encodes a Tag DNA polymerase enzyme
CC that has normal 5' nuclease activity, but reduced synthetic ability.
CC This sequence was derived from the mutag construct described in
CC AAT70343. The entire mutag gene was cut from the plasmid, and cloned
CC into pET-3c. This clone was digested with BstXI and BamHI, at unique
CC sites. The 3' overhang of BstXI was trimmed to a blunt end, while the 5'
CC overhang of BamHI was filled in. The blunt ends were ligated together.
CC This resulted in an in-frame deletion of 903 nucleotides. The enzyme
CC encoded by the present protein is also referred to as Cleavase BB.
XX
XX Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 0.0649 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
XX
XX US-09-823-645 3 (1-11) x AAT70347 (1-1600)
XX
QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
DB 1133 CTCCTCCAGAGGTAGCCATCCCTTACGAGGAG 1165
XX
XX AAT76647 standard; DNA; 1600 BP.
XX
XX AAT76647:
XX
XX 14-APR-1998 (first entry)
XX
XX Tag gene 5' nuclease clone 3F (Cleavase BB).
XX
XX Nucleic acid cleavage; DNA cleavage; RNA cleavage; 5' nuclease;
KM Tag; DNA polymerase; Cleavase BB; ds.
XX
XX Thermus aquaticus YT-1.
OS Synthetic.
XX
XX W09727214-A1.
PN
XX 31-JUL-1997.
PD
XX 22-JAN-1997; 97WO-US01072.
PF
XX 02-DEC-1996; 96US-0759038.
PR 24-JAN-1996; 96US-0599491.
XX 12-JUL-1996; 96US-0682853.
PR 29-NOV-1996; 96US-0756386.
XX 02-DEC-1996; 96US-0758314.
PR
XX

... (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX PA
XX Brow MAD, Dahlberg JE, Hall JG, Kaiser MM, Lyamichev VI;
PI Olive DM, Prudent JR;
XX WPI; 1997-393613/36. XX
XX
XX Thermostable structure-specific nuclease(s) - used for detection and
PT characterisation of nucleic acid sequences and variations in nucleic
PI acid sequences
XX
XX Example 2; Page 245; 457pp; English. XX
XX
XX This DNA sequence, denoted clone 3F or Cleavase BB, comprises an
CC altered Thermus aquaticus DNA polymerase (Tag) gene in which
CC nucleotides 875-1778 of the wild-type gene coding sequence are
CC deleted. Mutant gene mutTag (see AAT76643) was used as the starting
CC material for the construct. Cleavase BB is a thermostable
CC structure-specific nuclease preferred for use in nucleic acid
CC cleavage methods of the invention. Mutant genes (AAT76644-47)
CC were constructed in order to determine which portions of the Tag
CC polymerase domain can be altered without eliminating 5' nuclease
CC activity. The invention relates to means for the detection and
CC characterisation of nucleic acid (NA) sequences and variations in
CC NA. It relates to methods for forming a NA cleavage structure on
CC a target sequence and cleaving the NA cleavage structure in a
CC site-specific manner. The 5' nuclease activity of various enzymes
CC (see AAT4210-13) is used to cleave the target-dependent cleavage
CC structure, thereby indicating the presence of specific NA sequences
CC or specific variations of them.
XX
XX Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 0.0649 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
XX
XX US-09-823-649A-3 (1-11) x AAT76647 (1-1600)
XX
QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
DB 1133 CTCCTCCAGAGGTAGCCATCCCTTACGAGGAG 1165
XX
XX AAT76647 standard; DNA; 1600 BP.
XX
XX AAT76647:
XX
XX 02-FEB-1999 (first entry)
XX
XX Thermus aquaticus nuclease clone 3F (Cleavase BB) DNA.
XX
XX Nucleic acid detection; multiple sequential invasive cleavage;
KM DNA polymerase; nuclease; Cleavase BB; ds.
XX
XX Thermus aquaticus strain YT-1.
OS Synthetic.
XX
XX W09842873-A1.
PN
XX 01-OCT-1998.
PD
XX 24-MAR-1998; 98WO-US05809.
PF
XX 24-MAR-1997; 97US-0823516.
PR
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX


```
XX  Brow MAD, Ball JG, Kwiatkowski RM, Lyamichev VI;
PI  Mast AL, Vavra SH;
XX
XX  WPI; 1998-557036/47.
XX
XX  Detecting target nucleic acid by sequence-specific cleavage of
PI  complex with two specific oligonucleotides - used to detect
PI  cytomagalovirus DNA
XX
XX  Example 2; Page 278; 524pp; English.
XX
XX  This nucleotide sequence (clone 3f) codes for a thermostable
CC  nuclease (cleavase BB) derived from the DNA polymerase (DNAP) of
CC  Thermus aquaticus (Tag). In comparison to the wild-type Tag DNAP
CC  sequence (see AAV65779), it contains an in-frame deletion of 903
CC  nucleotides. The invention relates to means for the detection and
CC  characterisation of nucleic acid sequences, and variations in
CC  nucleic acid sequences. It also relates to methods for forming a
CC  nucleic acid cleavage structure on a target sequence and cleaving
CC  this structure in a site-specific manner, preferably using a
CC  thermostable structure-specific nuclease such as a modified Tag
CC  DNAP that has reduced synthetic activity (see AAV65783-86). Cleavage
CC  of the cleavage structure by the nuclease indicates the presence of
CC  specific nucleic acid sequences or specific variants. The invention
CC  further relates to methods for the separation of nucleic acid
CC  molecules based on charge, methods for the detection of non-target
CC  cleavage products via the formation of a complete and activated
CC  protein binding region, and methods for the detection of nucleic
CC  acid from various viruses (e.g. human cytomegalovirus) in a sample.
CC  The method amplifies the detection molecule rather than the target
CC  itself, is less subject to contamination than exponential
CC  amplification processes, and allows many targets to be analysed in
CC  a single reaction.
XX
XX  Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
SQ
XX
XX  Alignment Scores:
XX  Pred. No.: 0.0649 Length: 1600
XX  Score: 54.00 Matches: 11
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 100.00% Indels: 0
XX  DB: 19 Gaps: 0
XX
XX  US-09-823-649a-3 (1-11) x AAV65786 (1-1600)
OY  1 LeuSerGlnGluLeuAlaIleProTyrGlnGlu 11
DB  1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165
XX
XX  RESULT 6
XX  AAV53855
XX  AAV53855 standard; DNA; 1600 BP.
XX
XX  AAV53855;
XX
XX  21-DEC-1998 (first entry)
XX
XX  Nucleotide sequence of clone 4F of the Tag gene mutant.
XX
XX  Clone 4D: Tag mutant gene; thermostable; structure-specific nuclease;
KW  mutant DNA polymerase; bacteria; fungi; protozoa; RNA virus;
KW  hepatitis C virus; HCV; ds.
XX
XX  Thermus sp.
XX
XX  WO9823774-A1.
XX
XX  04-JUN-1998.
XX
XX  26-NOV-1997; 97WO-US21783.
XX
```

```
PR  02-DEC-1996; 96US-0758314.
PR  29-NOV-1996; 96US-0757653.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX  Kaiser MW, Lyamichev VI, Lyamicheva N;
XX
XX  WPI; 1998-322748/28.
XX
XX  Thermostable structure-specific nuclease(s) derived from mutant DNA
PI  polymerase(s) - useful for detecting mutant allele(s) or strains of
PI  microorganisms
XX
XX  Example 2; Page 238; 472pp; English.
XX
XX  This is the nucleotide sequence of a clone of a mutant Tag gene,
CC  used in the method of the invention. In this process thermostable
CC  structure-specific nucleases are derived from mutant DNA polymerases,
CC  which can be used for detecting mutant alleles or strains of
CC  microorganisms. The structure-specific nucleases can be used in
CC  mixtures, compositions and kits to treat nucleic acid, e.g. for
CC  detection of wild type and mutant alleles of genes, for detection
CC  and/or identification of strains of microorganisms such as bacteria,
CC  fungi, protozoa, especially for detection of RNA viruses such as the
CC  hepatitis C virus (HCV).
XX
XX  Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
SQ
XX
XX  Alignment Scores:
XX  Pred. No.: 0.0649 Length: 1600
XX  Score: 54.00 Matches: 11
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 100.00% Indels: 0
XX  DB: 19 Gaps: 0
XX
XX  US-09-823-649a-3 (1-11) x AAV53855 (1-1600)
OY  1 LeuSerGlnGluLeuAlaIleProTyrGlnGlu 11
DB  1133 CTCCTCCAGAGACTAGCCATCCCTTAGAGAGAG 1165
XX
XX  RESULT 7
XX  AAV63407
XX  AAV63407 standard; DNA; 1600 BP.
XX
XX  AAV63407;
XX
XX  26-JAN-1999 (first entry)
XX
XX  DNA sequence of a thermostable 5' nuclease derived from Tag polymerase.
XX
XX  Thermostable DNA polymerase; nucleic acid detection; cleavase BB;
KW  thermostable 5' nuclease; ds.
XX
XX  Synthetic.
OS  Thermus aquaticus.
XX
XX  US5837450-A.
XX
XX  17-NOV-1998.
XX
XX  06-JUN-1995; 95US-0471066.
XX
XX  06-JUN-1994; 94US-0254359.
XX  07-DEC-1992; 92US-0986330.
XX  04-JUN-1993; 93US-0073384.
XX  06-JUN-1995; 95US-0471066.
XX
XX  (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX  Brow MAD, Dahlberg JE, Lyamichev VI;
XX
```


DR WPI: 1999-023438/02.
 XX
 PT Detection of target nucleic acid molecules - uses modified
 PT thermostable enzymes with specific cleavage activity to create
 PT specific detection products from oligo:nucleotide(s) and target
 PT hybridisation
 XX
 PS Claim 10; Columns 79-82; 91pp; English.
 XX
 CC The present sequence represents a thermostable 5' nuclease derived from
 CC a thermostable polymerase modified to have reduced synthetic activity,
 CC where the 5' nuclease is capable of cleaving a linear nucleic acid
 CC duplex structure to create a single, single-stranded cleavage product.
 CC The nuclease, designated cleavage SB, is used in a method for detecting
 CC the presence of a nucleic acid molecule. The method is used for the
 CC specific detection of nucleic acid sequences, via a cleavage-based
 CC procedure, but without the need for amplification of target sequences.
 CC Thermostable polymerases, altered to have nuclease, but not polymerase
 CC activity are preferably used due to their specificity. The cleavage
 CC product specifically formed is detected, preferably by the use of
 CC radioactively labelled oligonucleotides. These can be used in
 CC e.g. forensic testing or paternity determination.
 CC
 SQ Sequence 1600 BP; 286 A; 519 G; 553 C; 242 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0649 Length: 1600
 Score: 54.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0
 US-09-823-649A-3 (1-11) x AAV63407 (1-1600)
 QY 1 leuSerGInGluLeuAlaIleProTyrgInGlu 11
 Db 1133 CTCTCCAGAGCTAGCCATCCTTAGCAGAG 1165
 RESULT 8
 ID AAQ23998 standard; DNA; 1635 BP.
 XX
 AC AAQ23998;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase pLSC8.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
 XX
 OS Thermus aquaticus.
 XX
 FH Key Location/Qualifiers
 FT old_sequence 1..2
 FT /*tag- a
 FT /note- "nucleotides 4-867 deleted from native
 XX sequence"
 XX
 PN WO9206200-A.
 PD 16-APR-1992.
 XX
 PD 16-APR-1992.
 XX
 PF 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelfand DH;

DR WPI: 1992-150885/18.
 DR P-PSDB; AAR23145.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC The sequence is that of Thermus aquaticus polymerase DNA which has
 CC been mutated. The mutation, resulting in mutant pLSC8, causes
 CC the polymerase enzyme produced to exhibit a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. Thermostable DNA
 CC polymerases are useful in many recombinant DNA techniques, esp.
 CC nucleic acid amplification by PCR, self-sustained sequence
 CC replication (SSR) and high temp. DNA sequencing. The absence of 5'-3'
 CC nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be
 CC desirable in enzymes used in homogeneous assays for the amplification
 CC and detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare a
 CC range of recombinant proteins having 5'-3' exonuclease activity to a
 CC complete lack of activity.
 CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
 XX
 SQ Sequence 1635 BP; 274 A; 551 C; 564 G; 246 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0666 Length: 1635
 Score: 54.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-823-649A-3 (1-11) x AAQ23998 (1-1635)
 QY 1 leuSerGInGluLeuAlaIleProTyrgInGlu 11
 Db 1168 CTCTCCAGAGCTAGCCATCCTTAGCAGAG 1200
 RESULT 9
 ID AAQ24322 standard; DNA; 1635 BP.
 XX
 AC AAQ24322;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase pL05A292.
 XX
 KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
 XX
 OS Thermus species Z05.
 XX
 FH Key Location/Qualifiers
 FT old_sequence 1..2
 FT /*tag- a
 FT /note- "nucleotides 4-873 deleted from the native
 XX sequence."
 XX
 PN WO9206200-A.
 PD 16-APR-1992.
 XX
 PD 16-APR-1992.
 XX
 PF 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.

```

XX  Abramson RD, Gelfand DH;
XX
XX  WPI: 1992-150885/18.
XX  P-PSDB: AAR23163.
XX
XX  Thermostable DNA polymerases with altered 5'-3' exo nuclease
XX  activity - having conserved regions mutated or deleted, for use
XX  in e.g. PCR, sequencing and detection assays
XX
XX  Claim 11, Page 59, 185pp; English.
XX
XX  The sequence is that of Thermus species Z05 polymerase DNA which
XX  has been mutated. The mutation designated p205A292 causes the
XX  polymerase enzyme produced to exhibit a different amt. of 5'-3'
XX  exonuclease activity than the native enzyme. Thermostable DNA
XX  polymerases are useful in many recombinant DNA techniques, esp.
XX  nucleic acid amplification by PCR, self-sustained sequence
XX  replication (SSR) and high temp. DNA sequencing. The absence of
XX  5'-3' nuclease activity may facilitate higher sensitivity allelic
XX  discrimination in a combined polymerase ligase chain reaction (PLCR)
XX  assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
XX  in enzymes used in homogeneous assays for the amplification and
XX  detection of a target nucleic acid sequence. Mutation of the DNA
XX  encoding particular regions of the enzymes can be used to prepare
XX  a range of recombinant proteins having 5'-3' exonuclease activity
XX  to a complete lack of activity.
XX  See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
XX  Sequence 1635 BP; 291 A; 577 C; 529 G; 238 T; 0 other;
XX
XX  Alignment Scores:
XX  Pred. No.: 0.0666 Length: 1635
XX  Score: 54.00 Matches: 11
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 100.00% Indels: 0
XX  DB: 13 Gaps: 0
XX
XX  US-09-823-649A-3 (1-11) x AAQ24322 (1-1635)
XX
XX  QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
XX  |||||||
XX  Db 1168 CTCCTCCAGAGCTTCCTCCATCCCTACGAGAG 1200
XX
XX  RESULT 10
XX  AAQ24328
XX  ID AAQ24328 standard; DNA: 1635 BP.
XX
XX  AC AAQ24328:
XX
XX  DT 22-Oct-1992 (first entry)
XX
XX  DE Mutant thermostable DNA polymerase pTHA292.
XX
XX  KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR; ss.
XX  OS Thermus thermophilus.
XX
XX  FH Key Location/Qualifiers
XX  FT old_sequence 1.2
XX  FT /*tag= a
XX  FT /note= "nucleotides 4-873 deleted from the native
XX  sequence."
XX
XX  MO9206200-A.
XX
XX  PD 16-APR-1992.
XX
XX  PF 30-SEP-1991; 91WO-US07035.
XX
XX  PR 28-SEP-1990; 90US-0590213.
XX  PR 28-SEP-1990; 90US-0590466.

```

```

PR 28-SEP-1990; 90US-0590490.
XX
XX  (CERTU ) CERTUS CORP.
XX
XX  Abramson RD, Gelfand DH;
XX
XX  WPI: 1992-150885/18.
XX  P-PSDB: AAR23168.
XX
XX  Thermostable DNA polymerases with altered 5'-3' exo nuclease
XX  activity - having conserved regions mutated or deleted, for use
XX  in e.g. PCR, sequencing and detection assays
XX
XX  Claim 11, Page 59, 185pp; English.
XX
XX  The sequence is that of Thermus thermophilus polymerase DNA which
XX  has been mutated. The mutation designated pTHA292 causes the
XX  polymerase enzyme produced to exhibit a different amt. of 5'-3'
XX  exonuclease activity than the native enzyme. Thermostable DNA
XX  polymerases are useful in many recombinant DNA techniques, esp.
XX  nucleic acid amplification by PCR, self-sustained sequence
XX  replication (SSR) and high temp. DNA sequencing. The absence of
XX  5'-3' nuclease activity may facilitate higher sensitivity allelic
XX  discrimination in a combined polymerase ligase chain reaction (PLCR)
XX  assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
XX  in enzymes used in homogeneous assays for the amplification and
XX  detection of a target nucleic acid sequence. Mutation of the DNA
XX  encoding particular regions of the enzymes can be used to prepare
XX  a range of recombinant proteins having 5'-3' exonuclease activity
XX  to a complete lack of activity.
XX  See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
XX  Sequence 1635 BP; 283 A; 577 C; 537 G; 238 T; 0 other;
XX
XX  Alignment Scores:
XX  Pred. No.: 0.0666 Length: 1635
XX  Score: 54.00 Matches: 11
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 100.00% Indels: 0
XX  DB: 13 Gaps: 0
XX
XX  US-09-823-649A-3 (1-11) x AAQ24328 (1-1635)
XX
XX  QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
XX  |||||||
XX  Db 1168 CTCCTCCAGAGCTTCCTCCATCCCTACGAGAG 1200
XX
XX  RESULT 11
XX  AAQ29460
XX  ID AAQ29460 standard; DNA: 1682 BP.
XX
XX  AC AAQ29460:
XX
XX  DT 14-MAR-2000 (first entry)
XX
XX  DE Thermostable DNA dependent mutant Tag polymerase-1 encoding DNA.
XX
XX  KW Thermostable DNA dependent polymerase; nuclease activity; Tag polymerase;
XX  mutant; Polymerase Chain Reaction; PCR; N-terminal domain; proteolytic;
XX  thermal stability; hydrophilic solution; sequencing; labeling; diagnosis;
XX  genetic disorder; primer-driven mutagenesis; identification of pathogen;
XX  mutational analysis; forensic identification; ds.
XX  OS Thermus aquaticus.
XX  OS Synthetic.
XX
XX  FH Key Location/Qualifiers
XX  FT CDS 8..1669
XX  FT /*tag= a
XX  FT /product= "Thermostable mutant Tag polymerase-1"
XX  FT /note= "identical at C-terminus to Tag polymerase
XX  from Thermus aquaticus"

```

```

XX PN WO964438-A1.
XX 16-DEC-1999.
XX 11-JUN-1999; 99WO-US13305.
XX 11-JUN-1998; 98US-0096399.
XX (CLON-) CLONTECH LAB INC.
XX Wurst H, Qui Z;
XX WPI; 2000-105869/09.
XX P-PSDB; AAY44352.
XX
XX PT Novel enzyme, particularly useful in polymerase chain reaction
XX
XX Claim 22; Fig 1; 32pp; English.
XX
XX The present sequence is the DNA encoding a novel, thermostable DNA
XX dependent, mutant Taq polymerase-1 enzyme. The N-terminal domain of the
XX polymerase includes a sequence of 9 amino acids, that has 40-50%
XX sequence identity with residues 280-288 of native Thermus aquaticus
XX polymerase, Tag. The residues 10-553 of encoded by this sequence is
XX identical to residues 289-832 of Taq polymerase. This sequence has no
XX significant nuclease activity, but good proteolytic and thermal
XX stabilities and improved solubility in hydrophilic solutions. This
XX sequence is used for synthesis of polynucleotides, particularly in
XX polymerase chain reaction (PCR) based processes, like sequencing,
XX labeling, primer-driven mutagenesis, diagnosis of genetic disorders,
XX identification of pathogens, mutational analysis and forensic
XX identification.
XX
SQ Sequence 1682 BP; 286 A; 567 C; 575 G; 254 T; 0 other;

Alignment Scores:
Pred. No.: 0.0689 Length: 1682
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-823-649A-3 (1-11) x AA229460 (1-1682)
QY 1 LeuSerGlnGluAlaIleProTyrGluGlu 11
Db 1202 CTCCTCCAGAGCTAGCCATCCCTTACGAGSAG 1234

RESULT 12
AAT47959
ID AAT47959 standard; DNA; 1686 BP.
XX
XX AAT47959;
XX
XX 11-MAY-1997 (first entry)
XX
XX Mutant Taq polymerase FY2.
XX
XX Tag; polymerase; FY2; FY3; FY4; thermostable; Thermus; enzyme;
XX DNA sequencing; PCR; ss.
XX
XX Thermus aquaticus.
XX
XX EP745676-A1.
XX
XX 04-DEC-1996.
XX
XX 30-MAY-1996; 96EP-0303880.
XX
XX 31-MAY-1995; 95US-0455686.
XX

```

```

PA (AMSH ) AMERSHAM LIFE SCI INC.
XX
XX PI Davis M, Fuller C, Moffett B, Fuller CM, Moffett RB;
XX WPI; 1997-013699/02.
XX P-PSDB; AAM09315.
XX
XX PT New enzymatically active Thermus DNA polymerase(s) - having a
XX tyrosine at a position equivalent to 667 of Tag DNA polymerase and
XX lacking 5' to 3' exonuclease activity
XX
XX Claim 5; Page 10-13; 36pp; English.
XX
XX FY2, FY3 and FY4 (see AAT47959 to AAT47961) are new thermostable DNA
XX polymerases providing improved results in DNA sequencing reactions.
XX FY2 is the T. aquaticus DNA polymerase lacking pre1. The 271-272
XX N-terminal amino acids and having the amino acid at position 667 of
XX the wild type (Phe) replaced by Tyr. FY3 contains 562 amino acids
XX with a Met at position 1 and Ala at position 2, corresponding to the
XX Met and Ala of positions 271 and 272, respectively, of the wild type
XX enzyme. FY4 corresponds to the T. thermophilus DNA polymerase
XX lacking the N-terminal 274 amino acids having a Phe to Tyr
XX replacement at position 396 (position 669 of the wild type enzyme).
XX
SQ Sequence 1686 BP; 283 A; 567 C; 578 G; 258 T; 0 other;

Alignment Scores:
Pred. No.: 0.0691 Length: 1686
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-823-649A-3 (1-11) x AAT47959 (1-1686)
QY 1 LeuSerGlnGluAlaIleProTyrGluGlu 11
Db 1219 CTCCTCCAGAGCTAGCCATCCCTTACGAGSAG 1251

RESULT 13
AAT47961
ID AAT47961 standard; DNA; 1686 BP.
XX
XX AAT47961;
XX
XX 11-MAY-1997 (first entry)
XX
XX Mutant Taq polymerase FY4.
XX
XX Tag; polymerase; FY2; FY3; FY4; thermostable; Thermus; enzyme;
XX DNA sequencing; PCR; ss.
XX
XX Thermus thermophilus.
XX
XX EP745676-A1.
XX
XX 04-DEC-1996.
XX
XX 30-MAY-1996; 96EP-0303880.
XX
XX 31-MAY-1995; 95US-0455686.
XX
XX (AMSH ) AMERSHAM LIFE SCI INC.
XX
XX Davis M, Fuller C, Moffett B, Fuller CM, Moffett RB;
XX WPI; 1997-013699/02.
XX P-PSDB; AAM09317.
XX
XX New enzymatically active Thermus DNA polymerase(s) - having a
XX tyrosine at a position equivalent to 667 of Tag DNA polymerase and
XX lacking 5' to 3' exonuclease activity
XX

```

XX Claim 5; Page 20-23; 36pp; English.

PS

CC FY2, FY3 and FY4 (see AAT47959 to AAT47961) are new thermostable DNA polymerases providing improved results in DNA sequencing reactions.

CC CC FY2 is the T. aquaticus DNA polymerase lacking pref. the 271-272

CC N-terminal amino acids and having the amino acid at position 667 of

CC the wild type (Phe) replaced by Tyr. FY3 contains 562 amino acids

CC with a Met at position 1 and Ala at position 2, corresponding to the

CC Met and Ala of positions 271 and 272, respectively, of the wild type

CC enzyme. FY4 corresponds to the T. thermophilus DNA polymerase

CC lacking the N-terminal 274 amino acids having a Phe to Tyr

CC replacement at position 396 (position 669 of the wild type enzyme).

XX

SQ Sequence 1686 BP; 295 A; 593 C; 551 G; 247 T; 0 other;

Alignment Scores:

Pred. No.:	0.0691	Length:	1686
Score:	54.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	18	Gaps:	0

US-09-823-649a-3 (1-11) x AAT47961 (1-1686)

QY 1 LeuSerGlnIuLeuAlaIleProTyrGluGlu 11

DB 1219 CTCGCCAGCAGCTACCATCCCTCCAGAGAA 1251

RESULT 14

AAAX27130

ID AAX27130 standard; DNA; 1686 BP.

XX

AC AAX27130;

XX

DT 27-MAY-1999 (first entry)

XX

DE FY2 polymerase coding sequence.

XX

KW FY2 polymerase; FY3 polymerase; FY4 polymerase; Taq polymerase variant;

XX

KW DNA sequencing; ss.

XX

OS Thermus aquaticus.

XX

PN US585813-A.

XX

PD 23-MAR-1999.

XX

PF 14-MAY-1996; 96US-0648657.

XX

PR 14-MAY-1996; 96US-0648657.

XX

PR 31-MAY-1995; 95US-0455686.

XX

PA (AMSH) AMERSHAM LIFE SCI INC.

XX

PI Cunniff JJ, Davis M, Fuller CW, Moffett RB;

XX

DR WPI; 1999-228538/19.

XX

DR P-PSDB; AAY00886.

XX

PT New thermostable DNA polymerases - having 540 to 582 amino acids

XX

PT and lacking 5' to 3' exonuclease activity

XX

PS Claim 7; Column 11-16; 27pp; English.

XX

CC This sequence encodes a thermostable DNA polymerase of the invention. The

CC thermostable polymerases of the invention are variants of Taq DNA

CC polymerase having a Tyr residue at a position corresponding to Taq DNA

CC polymerase residue 667 in its dNMP binding site, and lacking 5' to 3'

CC exonuclease activity as a result of an N-terminal deletion. The

CC polymerases can be used for the sequencing of DNA, either by manual or

CC automated means.

XX

SQ Sequence 1686 BP; 283 A; 567 C; 578 G; 258 T; 0 other;

Alignment Scores:

Pred. No.:	0.0691	Length:	1686
Score:	54.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-09-823-649a-3 (1-11) x AAX27130 (1-1686)

QY 1 LeuSerGlnIuLeuAlaIleProTyrGluGlu 11

DB 1219 CTCGCCAGCAGCTACCATCCCTTACGAGAG 1251

RESULT 15

AAAX27132

ID AAX27132 standard; DNA; 1686 BP.

XX

AC AAX27132;

XX

DT 27-MAY-1999 (first entry)

XX

DE FY4 polymerase coding sequence.

XX

KW FY2 polymerase; FY3 polymerase; FY4 polymerase; Taq polymerase variant;

XX

KW DNA sequencing; ss.

XX

OS Thermus thermophilus.

XX

PN US585813-A.

XX

PD 23-MAR-1999.

XX

PF 14-MAY-1996; 96US-0648657.

XX

PR 14-MAY-1996; 96US-0648657.

XX

PR 31-MAY-1995; 95US-0455686.

XX

PA (AMSH) AMERSHAM LIFE SCI INC.

XX

PI Cunniff JJ, Davis M, Fuller CW, Moffett RB;

XX

DR WPI; 1999-228538/19.

XX

DR P-PSDB; AAY00888.

XX

PT New thermostable DNA polymerases - having 540 to 582 amino acids

XX

PT and lacking 5' to 3' exonuclease activity

XX

PS Claim 7; Column 23-28; 27pp; English.

XX

CC This sequence encodes a thermostable DNA polymerase of the invention. The

CC thermostable polymerases of the invention are variants of Taq DNA

CC polymerase having a Tyr residue at a position corresponding to Taq DNA

CC polymerase residue 667 in its dNMP binding site, and lacking 5' to 3'

CC exonuclease activity as a result of an N-terminal deletion. The

CC polymerases can be used for the sequencing of DNA, either by manual or

CC automated means.

XX

SQ Sequence 1686 BP; 295 A; 593 C; 551 G; 247 T; 0 other;

Alignment Scores:

Pred. No.:	0.0691	Length:	1686
Score:	54.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-09-823-649a-3 (1-11) x AAX27132 (1-1686)

Qy 1 LeuSerGlnGluLeuAlaIleProTyrGlu 11
|||||
Db 1219 CTCTCCGAGAGCTAGCCATCCCTACGAGAA 1251

Search completed: January 15, 2003, 12:55:28
Job time : 131.143 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:02:09 ; Search time 28.1429 Seconds

(without alignments)
52.083 Million cell updates/sec

Title: US-09-823-649A-3

Sequence: 1 LSOEALPYEE 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq.101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length DB	ID	Description
1	54	100.0	11	AA847793	Native DNA polymer
2	54	100.0	11	AA848259	Native DNA polymer
3	54	100.0	11	AA848261	Native DNA polymer
4	54	100.0	11	AA848262	Native DNA polymer
5	54	100.0	11	AA848264	Native DNA polymer
6	54	100.0	17	AA896267	Mutant Thermus aqu
7	54	100.0	544	AA823145	Mutant Thermostabl
8	54	100.0	544	AA823163	Mutant Thermostabl
9	54	100.0	544	AA823168	Mutant Thermostabl
10	54	100.0	552	AA809304	Thermus aquaticus

11	54	100.0	553	21	AA14352	Thermostable DNA d
12	54	100.0	554	16	AA86209	Novel thermostable
13	54	100.0	559	21	AA14353	Thermostable DNA d
14	54	100.0	560	17	AA899544	Thermus flavus DNA
15	54	100.0	560	17	AA896205	DNA polymerase I e
16	54	100.0	561	18	AA809315	Mutant Taq polymer
17	54	100.0	561	18	AA809317	Mutant Taq polymer
18	54	100.0	561	20	AA100888	Fy2 polymerase. T
19	54	100.0	562	18	AA100888	Fy4 polymerase. T
20	54	100.0	562	20	AA100887	Mutant Taq polymer
21	54	100.0	562	20	AA100887	Fy3 polymerase. T
22	54	100.0	562	22	AA809305	Thermus thermophil
23	54	100.0	562	22	AA809310	Thermus aquaticus
24	54	100.0	597	17	AA899543	Thermus flavus DNA
25	54	100.0	597	17	AA896204	DNA polymerase I e
26	54	100.0	604	22	AA800574	Chimeric Cautaq DN
27	54	100.0	604	22	AA800575	Chimeric Cautaq DN
28	54	100.0	631	22	AA800575	Mutant thermostabl
29	54	100.0	632	13	AA823144	Mutant thermostabl
30	54	100.0	632	13	AA823162	Mutant thermostabl
31	54	100.0	632	13	AA823167	Mutant thermostabl
32	54	100.0	632	23	AA815565	Sulfolobus solifata
33	54	100.0	654	23	AA815569	Sulfolobus acidoca
34	54	100.0	680	13	AA823143	Polylysine (PL) -
35	54	100.0	680	13	AA823161	Mutant thermostabl
36	54	100.0	680	13	AA823161	Mutant thermostabl
37	54	100.0	757	13	AA823141	Mutant thermostabl
38	54	100.0	758	13	AA823160	Mutant thermostabl
39	54	100.0	758	13	AA823165	Mutant thermostabl
40	54	100.0	787	13	AA823142	Mutant thermostabl
41	54	100.0	789	13	AA823159	Mutant thermostabl
42	54	100.0	789	13	AA823164	Mutant thermostabl
43	54	100.0	810	22	AA862314	T. aquaticus (Taq)
44	54	100.0	810	23	ABB08339	Wild type Taq DNA
45	54	100.0	810	23	ABB08340	Cold-sensitive mut

ALIGNMENTS

RESULT 1	AA847793	standard: peptide; 11 AA.
ID	AA847793	
AC	AA847793	
DT	25-MAR-2002	(first entry)
DE	Native DNA polymerase motif #3.	
XX		
KW	DNA polymerase; reverse transcription; primer; divalent cation; mutant;	
KW	transverse transcription reaction; fluorescein; cyanine; thermocactive;	
KW	dye; amplification.	
XX		
OS	Thermus sp.	
XX		
PN	EP1152062-A2.	
XX		
PD	07-NOV-2001.	
XX		
PF	12-APR-2001; 2001EP-0109341.	
XX		
PR	18-APR-2000; 2000US-198336P.	
XX		
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX		
PI	Smith ES, Elfstrom CW, Gelfand DH, Higuchi RG, Myers TW;	
PI	Schoenrunner NJ, Wang AM;	
XX		
DR	WPI; 2002-076891/11.	
XX		
PT	Reverse transcribing an RNA, comprises performing a reverse	
PT	transcriptase polymerase chain reaction amplification of a mixture	

CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
XXXXXXXXXXXX
Db 1 LSOELAIPIYEE 11

RESULT 4
AAM48262
ID AAM48262 standard; Peptide: 11 AA.

AC AAM48262;
XX
DT 25-MAR-2002 (first entry)
XX
DE Native DNA polymerase motif #11.

XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KM transverse transcription reaction; fluorescein; cyanine; thermoactive;
KW dye; amplification.

XX Thermus species 205.

PN BP1152062-A2.

PD 07-NOV-2001.

PF 12-APR-2001; 2001EP-0109341.

PR 18-APR-2000; 2000US-198336P.

PA (HOFF) HOFMANN LA ROCHE & CO AG F.

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermoactive DNA polymerase -

PS Disclosure: Page 7; 23pp; English.

XX The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
XXXXXXXXXXXX
Db 1 LSOELAIPIYEE 11

RESULT 5
AAM48264
ID AAM48264 standard; Peptide: 11 AA.

AC AAM48264;
XX
DT 25-MAR-2002 (first entry)
XX
DE Native DNA polymerase motif #13.

XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KM transverse transcription reaction; fluorescein; cyanine; thermoactive;
KW dye; amplification.

XX Thermus caldophilus.

PN BP1152062-A2.

PD 07-NOV-2001.

PF 12-APR-2001; 2001EP-0109341.

PR 18-APR-2000; 2000US-198336P.

PA (HOFF) HOFMANN LA ROCHE & CO AG F.

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermoactive DNA polymerase -

PS Disclosure: Page 7; 23pp; English.

XX The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSOELAIPIYEE 11
XXXXXXXXXXXX
Db 1 LSOELAIPIYEE 11

RESULT 6
AAR96267
ID AAR96267 standard; Protein: 528 AA.

AC AAR96267;
 XX
 DT 18-NOV-1996 (first entry)
 XX
 DE Mutant *Thermus aquaticus* DNA polymerase (Clone 4F).
 XX
 DE p53; mutant; mutation; cleavage; nuclease; cleavage; *Thermus*;
 KW *Escherichia*; *Saccharomyces*; *Campylobacter*; *Mycobacterium*; *Shigella*;
 KM *Staphylococcus*; identification; detection.
 XX
 OS *Thermus aquaticus*.
 XX
 PN WO9615267-A1.
 XX
 PU 23-MAY-1996.
 XX
 PF 09-NOV-1995; 95WO-US14673.
 XX
 PR 30-AUG-1995; 95US-0520946.
 PR 09-NOV-1994; 94US-0337164.
 PR 09-MAR-1995; 95US-0402601.
 PR 07-JUN-1995; 95US-0484956.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Brow MAD, Dahlberg JE, Fors L, Heisler LM, Lyamichev VI;
 PI Oldenburg MC, Olive DM;
 XX
 DR WPI: 1996-259862/26.
 DR N-PSDB; AAT27686.
 XX
 PT Cleavage of nucleic acids to detect mutation(s) - allows detection
 PT esp. in human p53 gene, to identify strains of microorganisms and
 PT viruses
 XX
 PS Example 2a; Page 287-288; 433pp; English.
 XX
 CC Cleavage of nucleic acids using an enzyme, especially a nuclease
 CC selected from the group consisting of Cleavage (RTM) BN enzyme,
 CC *Thermus aquaticus* DNA polymerase, *Thermus thermophilus* DNA
 CC polymerase, *Escherichia coli* ExoIII and the *Saccharomyces cerevisiae*
 CC Rad1/Rad10 complex. The nucleic acid substrate is preferably an
 CC oligonucleotide containing a human p53 gene sequence or
 CC alternatively, microbial gene sequences. Cleavage products are
 CC compared to the cleavage products of reference gene sequences. The
 CC method is used for detecting mutation in the human p53 gene; for
 CC identifying strains of microorganisms, especially bacteria selected
 CC from the group of members of the genera *Campylobacter*,
 CC *Escherichia*, *Mycobacterium*, *Salmonella*, *Shigella* and *Staphylococcus*.
 CC The method may also be used for the identification of viruses,
 CC especially hepatitis C virus and simian immunodeficiency virus.
 CC *Thermus aquaticus* (Tag) DNA polymerase was amplified using two
 CC primers (AAT27679, AAT27680). The Tag polymerase DNA was inserted into
 CC the BamHI restriction site of the expression vector pET-3c and mutant
 CC genes were created from that construct. This mutant was created
 CC after the vector was digested with BstXI and BamHI. The DNA was
 CC then treated with the Klenow fragment of DNAPPC to trim both 3'
 CC overhangs to blunt ends which were then ligated together, resulting
 CC in an in frame deletion of 903 nucleotides. This mutant Tag
 CC polymerase is also referred to as the Cleavage BB enzyme.
 XX
 SQ Sequence 528 AA;
 XX

Query Match 100.0%; Score 54; DB 17; Length 528;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LSQELAIPEE 11
 DB 374 LSQELAIPEE 384

RESULT 7

AAR23145
 ID AAR23145 standard; Protein; 544 AA.
 XX
 AC AAR23145;
 XX
 DT 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase enzyme MET-SER 290 Tag.
 DE
 DE 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR.
 KW
 XX
 OS *Thermus aquaticus*.
 XX
 PN WO9206200-A.
 XX
 PU 16-APR-1992.
 XX
 PF 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelland DJ;
 PI
 XX
 DR WPI: 1992-150885/18.
 DR N-PSDB; AAQ23998.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exonuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC The sequence is that of a mutant of *Thermus aquaticus* polymerase
 CC mutant MET-SER 290 Tag (the Stoitel fragment) having a different
 CC amt. of 5'-3' exonuclease activity than the native enzyme.
 CC Thermostable DNA polymerases are useful in many recombinant DNA
 CC techniques, esp. nucleic acid amplification by PCR, self-sustained
 CC sequence replication (SSR) and high temp. DNA sequencing. The absence
 CC of 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAR23140-79 and AAR231722.
 XX
 SQ Sequence 544 AA;
 XX

Query Match 100.0%; Score 54; DB 13; Length 544;
 Best Local Similarity 100.0%; Pred. No. 0.076;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LSQELAIPEE 11
 DB 390 LSQELAIPEE 400

RESULT 8
 AAR23163
 ID AAR23163 standard; Protein; 544 AA.
 XX
 AC AAR23163;
 XX

```

XX DT 22-OCT-1992 (first entry)
XX DE Mutant thermostable DNA polymerase enzyme MET-ALA 292 T205.
XX DE 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
XX OS Thermus species 205.
XX PN WO9206200-A.
XX PD 16-APR-1992.
XX PF 30-SEP-1991; 91WO-US07035.
XX PR 28-SEP-1990; 90US-0590213.
XX PR 28-SEP-1990; 90US-0590466.
XX PR 28-SEP-1990; 90US-0590490.
XX PA (CETU ) CETUS CORP.
XX PI Abramson RD, Gelfand DH;
XX DR WPI: 1992-150885/18.
XX DR N-PSDB; AAQ24322.
XX PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
XX PT activity - having conserved regions mutated or deleted, for use
XX PT in e.g. PCR, sequencing and detection assays
XX PS Claim 11; Page 59; 185pp; English.
XX CC The sequence is that of a mutant of Thermus species 205 polymerase
XX CC mutant MET-ALA 292 T205, having a different amt. of 5'-3' exonuclease
XX CC activity than the native enzyme. Thermostable DNA polymerases
XX CC are useful in many recombinant DNA techniques, esp. nucleic acid
XX CC amplification by PCR, self-sustained sequence replication (SSR)
XX CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
XX CC may facilitate higher sensitivity allelic discrimination in a combined
XX CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
XX CC 5'-3' exonuclease activity may be desirable in enzymes used in
XX CC homogeneous assays for the amplification and detection of a target
XX CC nucleic acid sequence. Mutation of the DNA encoding particular
XX CC regions of the enzymes can be used to prepare a range of recombinant
XX CC proteins having 5'-3' exonuclease activity to a complete lack of
XX CC activity.
XX CC See also AAR3140-79 and AAR3722.
XX CS
XX SO Sequence 544 AA;

Query Match 100.0%; Score 54; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
DB 390 LSQELAIPIYEE 400

RESULT 9
AAR33168
ID AAR33168 standard; Protein; 544 AA.
XX AC AAR33168;
XX DT 22-OCT-1992 (first entry)
XX DE Mutant thermostable DNA polymerase enzyme MET-ALA 292 Tth.
XX DE 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
XX OS Thermus thermophilus.
XX PN
XX PD

```

```

PH Key Location/Qualifiers
PH FT Misc-difference 1..2
PH FT "residues 2-291 deleted from the native
PH FT "sequence"
XX PN WO9206200-A.
XX PD 16-APR-1992.
XX PF 30-SEP-1991; 91WO-US07035.
XX PR 28-SEP-1990; 90US-0590213.
XX PR 28-SEP-1990; 90US-0590466.
XX PR 28-SEP-1990; 90US-0590490.
XX PA (CETU ) CETUS CORP.
XX PI Abramson RD, Gelfand DH;
XX DR WPI: 1992-150885/18.
XX DR N-PSDB; AAQ24328.
XX PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
XX PT activity - having conserved regions mutated or deleted, for use
XX PT in e.g. PCR, sequencing and detection assays
XX PS Claim 11; Page 59; 185pp; English.
XX CC The sequence is that of a mutant of Thermus thermophilus polymerase
XX CC mutant MET-ALA 292 Tth, having a different amt. of 5'-3' exonuclease
XX CC activity than the native enzyme. Thermostable DNA polymerases
XX CC are useful in many recombinant DNA techniques, esp. nucleic acid
XX CC amplification by PCR, self-sustained sequence replication (SSR)
XX CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
XX CC may facilitate higher sensitivity allelic discrimination in a combined
XX CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
XX CC 5'-3' exonuclease activity may be desirable in enzymes used in
XX CC homogeneous assays for the amplification and detection of a target
XX CC nucleic acid sequence. Mutation of the DNA encoding particular
XX CC regions of the enzymes can be used to prepare a range of recombinant
XX CC proteins having 5'-3' exonuclease activity to a complete lack of
XX CC activity.
XX CC See also AAR3140-79 and AAR3722.
XX CS
XX SO Sequence 544 AA;

Query Match 100.0%; Score 54; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPIYEE 11
DB 390 LSQELAIPIYEE 400

RESULT 10
AAE09304
ID AAE09304 standard; Protein; 552 AA.
XX AC AAE09304;
XX DT 22-NOV-2001 (first entry)
XX DE Thermus aquaticus (Tag) DNA polymerase #1.
XX DE Thermus aquaticus; Tag; DNA polymerase; polymerisation; PCR protocol;
XX DE genetic engineering.
XX OS Thermus aquaticus.
XX PN WO200161015-A2.
XX PD 23-AUG-2001.

```

XX PF 16-FEB-2001; 2001WO-EP01790.
 XX XX
 XX PR 17-FEB-2000; 2000US-0506153.
 XX XX
 XX PA (QIAGEN-) QIAGEN GMBH.
 XX PI Loeffert D, Missel A, Kang J;
 XX DR WPI; 2001-536571/59.
 XX XX
 PT Chimeric nucleic acid polymerase for nucleic acid polymerization and
 PT generation of nucleic acids, comprises two enzymatically active domains
 PT which are non-naturally associated
 XX
 XX PS Claim 12; Page 46-49; 93pp; English.
 XX CC The invention relates to a chimeric nucleic acid polymerase comprising
 CC at least two enzymatically active domains which are non-naturally
 CC associated. The polymerase is thermostable and possesses new or improved
 CC catalytic properties compared to known nucleic acid polymerases. The
 CC polymerase eliminates the need to specifically develop multi-enzyme
 CC reaction mixtures, which are often difficult to optimise and expensive
 CC to use. The polymerase facilitates rapid, efficient and accurate
 CC generation of nucleic acid molecules, particularly in regard to PCR
 CC protocols. The polymerase is also useful for nucleic acid polymerisation
 CC which is useful in genetic engineering techniques and molecular biology.
 CC The present protein sequence is *Thermus aquaticus* (Tag) DNA polymerase.
 XX
 SQ Sequence 552 AA:
 Query Match 100.0%; Score 54; DB 22; Length 552;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LSQELAIPEE 11
 |||||||||
 DB 398 LSQELAIPEE 408

RESULT 11
 ID AAY44352 standard; Protein: 553 AA.
 XX
 AC AAY44352;
 XX
 DT 14-MAR-1000 (first entry)
 XX
 DE Thermostable DNA dependent mutant Tag polymerase-1.
 XX
 KW Thermostable DNA dependent polymerase; nuclease activity; "tag polymerase;
 KW mutant; polymerase chain reaction; PCR; N-terminal domain; protocolytic;
 KW thermal stability; hydrophilic solution; sequencing; labeling; diagnosis;
 KW genetic disorder; primer-driven mutagenesis; identification of pathogen;
 KW mutational analysis; forensic identification.
 XX
 OS *Thermus aquaticus*.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FH Domain 1..9
 FT /label= "Nine residue domain
 FT /note= "40-50% identical to residues 280-288 of native
 FT *Thermus aquaticus*"
 FT 10..553
 FT /note= "Identical to residues 289-832 of native
 FT *Thermus aquaticus*"
 XX
 PN W09964438-A1.
 XX
 PD 16-DEC-1999.
 XX
 PF 11-JUN-1999; 99WO-US13305.

XX XX
 XX PR 11-JUN-1998; 98US-0096399.
 XX XX
 XX PA (CLON-) CLONTECH LAB INC.
 XX XX
 XX PI Wurst H, Qui Z;
 XX DR WPI; 2000-105869/09.
 XX DR N-PSDB; AA229460.
 XX
 PT Novel enzyme, particularly useful in polymerase chain reaction
 PS
 PS Claim 16; Fig 2; 32pp; English.
 XX
 CC The present amino acid sequence is a novel, thermostable DNA dependent
 CC mutant Tag polymerase-1 enzyme. The N-terminal domain, includes a
 CC sequence of 9 amino acids, that has 40-50% sequence identity with
 CC residues 280-288 of native *Thermus aquaticus* polymerase, Tag. The
 CC residues 10-553 of this polymerase sequence is identical to residues
 CC 289-832 of Tag polymerase. This sequence has no significant nuclease
 CC activity, but good protocolytic and thermal stabilities and improved
 CC solubility in hydrophilic solutions. This sequence is used for synthesis
 CC of polynucleotides, particularly in polymerase chain reaction (PCR),
 CC based processes, like sequencing, labeling, primer-driven mutagenesis,
 CC diagnosis of genetic disorders, identification of pathogens, mutational
 CC analysis and forensic identification.
 XX
 SQ Sequence 553 AA:
 Query Match 100.0%; Score 54; DB 21; Length 553;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LSQELAIPEE 11
 |||||||||
 DB 399 LSQELAIPEE 409

RESULT 12
 ID AAR66209 standard; Protein: 554 AA.
 XX
 AC AAR66209;
 XX
 DT 10-AUG-1995 (first entry)
 XX
 DE Novel thermostable DNA polymerase KlenTaq-278.
 XX
 KW Thermostable polymerase; KlenTaq 278.
 XX
 OS Synthetic.
 OS
 PN W09426766-A.
 PD 24-NOV-1994.
 XX
 PF 22-FEB-1994; 94WO-US01867.
 XX
 PR 19-FEB-1993; 93US-0021623.
 PR 22-FEB-1994; 94US-0021623.
 XX
 PA (BARN/) BARNES W M.
 XX
 PI Barnes WM;
 XX
 DR WPI; 1995-006692/01.
 DR N-PSDB; AA079545.
 XX
 PT DNA polymerase and formulations comprising it - allowing the
 PT amplification of sequences up to 35 kilobases and reducing the
 PT mutagenicity generated by the PCR process.
 XX
 PS Claim 3; Page 39-44; 79pp; English.

XX KlenTag-278 is a thermostable polymerase lacking 3'-exonuclease
 CC activity. The AA sequence is substantially the same AA sequence as
 CC Thermus aquaticus DNA polymerase but lacks the N-terminal 280 AAs. A
 CC DNA encoding such a polymerase is claimed, as is the polymerase
 CC itself. The polymerase also has substantially the same AA sequence
 CC as that of Thermus flavus, but excludes the N-terminal 279 AAs.
 CC Primers AA079543, AA079544, AA079553 and AA079554 can be used for the
 CC amplification of the gene for KlenTag-278. Essentially the same
 CC primers can be used for the amplification of the analogous gene from
 CC Thermus flavus (see AA079543, AA079556, AA079557). An initiator Met and a
 CC Gly occupy the first two N-terminal posns of KlenTag-278.
 CC previously occupied by residues 279 and 280 of T. aquaticus
 CC DNA polymerase, followed by the AA sequence of wt T. aquaticus
 CC DNA polymerase, beginning with the AA residue at posn. 281.
 CC Plasmid pMB254b contains the sequence in AA079545.

XX Sequence 554 AA;

Query Match 100.0%; Score 54; DB 16; Length 554;

Best local Similarity 100.0%; Pred. No. 0.078;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSGELAIPEE 11
 DB 400 LSGELAIPEE 410

RESULT 13

AAV44353
 ID AAV44353 standard; Protein; 559 AA.

XX AAV44353;

DT 14-MAR-2000 (first entry)

DE Thermostable DNA dependent mutant Taq polymerase-2.

XX Thermotable DNA dependent polymerase; nuclease activity; Taq polymerase;
 KW mutant; Polymerase Chain Reaction; PCR; N-terminal domain; proteolytic;
 KW thermal stability; hydrophilic solution; sequencing; labeling; diagnosis;
 KW genetic disorder; primer-driven mutagenesis; identification of pathogen;
 KW mutational analysis; forensic identification.

XX Thermus aquaticus.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Domain 1.6

FT /label= Leader-domain
 FT /note= "40-50% identical to N-terminal domain of native
 Thermus aquaticus"

FT Domain 7..15

FT /label= Nine residue domain
 FT /note= "40-50% identical with residues 280-288 of native
 Thermus aquaticus"

FT Region 16..559

FT /note= "identical to residues 289-832 of native
 Thermus aquaticus"

XX W09964438-A1.

XX 16-DEC-1999.

XX 11-JUN-1999; 99WO-US13305.

XX 11-JUN-1998; 98US-0096399.

XX (CLON-) CLONTECH LAB INC.

XX Wurst H, Qui Z;

XX MPI; 2000-105869/09.

DR N-PSDB; AA229461.

PT Novel enzyme, particularly useful in polymerase chain reaction

PS Claim 20; Fig 4; 32pp; English.

XX The present amino acid sequence is a novel, thermostable DNA dependent
 CC mutant Taq polymerase-2 enzyme. The N-terminal domain, includes a
 CC sequence of 15 amino acids, that has 40-50% sequence identity
 CC with N-terminal residues of native Thermus aquaticus polymerase. Tag
 CC 289-832 of Taq polymerase. This sequence has no significant nuclease
 CC activity, but good proteolytic and thermal stabilities and improved
 CC solubility in hydrophilic solutions. This sequence is used for synthesis
 CC of polynucleotides, particularly in polymerase chain reaction (PCR)
 CC based processes, like sequencing, labeling, primer-driven mutagenesis,
 CC diagnosis of genetic disorders, identification of pathogens, mutational
 CC analysis and forensic identification.

XX Sequence 559 AA;

Query Match 100.0%; Score 54; DB 21; Length 559;

Best local Similarity 100.0%; Pred. No. 0.079;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSGELAIPEE 11
 DB 405 LSGELAIPEE 415

RESULT 14

AA99544
 ID AA99544 standard; Protein; 560 AA.

XX AA99544;

DT 29-NOV-1996 (first entry)

DE Thermus flavus DNA polymerase I exonuclease free fragment.

XX DNA polymerase; thermostable; sequencing; amplification; PCR;
 KW polymerase chain reaction; universal thermal cycle labelling; UPCR;
 KW TCI; Ligase chain reaction.

XX Thermus flavus.

XX W09614417-A1.

XX 17-MAY-1996.

XX 03-NOV-1995; 95WO-US14418.

XX 04-NOV-1994; 94US-0334640.

XX (MOLE-) MOLECULAR BIOLOGY RESOURCES INC.

XX Mueller RD, Nickerson DA, Piehl RF, Skowron PW;

XX Swaminathan N;

XX MPI; 1996-251767/25.

XX DNA encoding thermostable Thermus flavus DNA polymerase - useful for
 PT DNA sequencing, polymerase chain reaction, thermal cycle labelling
 and ligase chain reaction, etc.

XX Disclosure; Page 93-95; 139pp; English.

XX DNA encoding a thermostable polymerase can be inserted into an
 CC expression vector which in turn can be used to transform cells. The
 CC transformed cells can then be cultured and the thermostable
 CC polymerase enzyme retrieved from the culture supernatant. Such
 CC recombinantly produced polymerase can be used in applications such
 CC as DNA sequencing, DNA amplification, thermal cycle labelling (TCI),

Search completed: January 15, 2003, 11:20:14
 Job time : 28.2857 secs

CC universal thermal cycle labelling (UTCL) and ligase chain reaction
 CC (LCK). This fragment of the exonuclease free fragment of DNA
 CC polymerase I retains its polymerase activity.
 XX

SQ Sequence 560 AA:

Query Match

100.0%; Score 54; DB 17; Length 560;
 Best Local Similarity 100.0%; Pred. No. 0.079;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSQELAIPEE 11
 |||||

DB 406 LSQELAIPEE 416

RESULT 15

AA96205

ID AA96205 standard; Protein; 560 AA.

XX AA96205;

XX

DT 29-NOV '95 (first entry)

XX

DE DNA polymerase I exonuclease-free fragment.

XX

KM DNA polymerase I; truncated; exonuclease free; holoenzyme; universal;

KW DNA sequencing; amplification; polymerase chain reaction;

XX ligase chain reaction; thermal cycle labelling.

XX Thermus flavus.

OS

PN WO9614405-A2.

XX

PD 17-MAY-1996.

XX

PF 03-NOV-1995; 95WO-US15327.

XX

PR 04-NOV-1994; 94US-0334645.

XX

PA (MOLE-) MOLECULAR BIOLOGY RESOURCES INC.

XX

PI Mueller RD, Piehl RF, Skowron PM, Swaminathan N;

XX

DR WP1; 1996-251756/25.

DR N-PSDB; AAT27255.

XX

PT

PT Biologically active fragments of Thermus flavus DNA polymerase
 useful for DNA sequencing, polymerase chain reaction, thermal cycle
 labelling and ligase chain reaction, etc.

XX

PS Claim 1; Page 93-95; 139pp; English.

XX

CC The present sequence is that of a truncated Thermus flavus DNA

CC polymerase I, which is exonuclease-free, and corresponds to amino acids

CC 275 to 834 of AA96203. A vector lacking the 5' one-third of the T.

CC flavus DNA polymerase I was generated. The ATG startcodon of lacZ was

CC brought in frame with the DNA encoding amino acids 239 to 834 of the Tfl

CC DNA pol I holoenzyme using site-directed mutagenesis. This was used to

CC generate the present sequence. This recombinant polymerase is

CC thermostable and can be used in applications such as DNA sequencing,

CC polymerase chain reaction, (universal) thermal cycle labelling and ligase

CC chain reaction.

XX

SQ Sequence 560 AA:

Query Match

100.0%; Score 54; DB 17; Length 560;
 Best Local Similarity 100.0%; Pred. No. 0.079;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSQELAIPEE 11
 |||||

DB 406 LSQELAIPEE 416

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 : Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec

Title: US-09-823-649a-4
Perfect score: 50
Sequence: 1 ISXELSTPYEE 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/CTUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	96.0	830	1	US-07-977-434-6
2	48	96.0	830	1	US-08-458-819-6
3	48	96.0	830	5	PCT-US91-07035-6
4	48	96.0	831	1	US-08-073-384C-5
5	48	96.0	831	1	US-08-254-359A-5
6	48	96.0	831	1	US-08-483-043-5
7	48	96.0	831	1	US-08-481-238-5
8	48	96.0	831	2	US-08-471-066B-5
9	48	96.0	831	2	US-08-484-956-5
10	48	96.0	831	2	US-08-757-653-5
11	48	96.0	831	2	US-08-599-491-5
12	48	96.0	831	2	US-08-756-386-5
13	48	96.0	831	2	US-08-823-516-5
14	48	96.0	831	2	US-08-682-853A-5
15	48	96.0	831	3	US-08-759-038-5
16	48	96.0	831	3	US-08-758-314-5
17	48	96.0	831	4	US-09-350-309-5
18	48	96.0	831	4	US-08-520-946-5
19	48	96.0	831	4	US-08-484-956-80
20	48	96.0	831	2	US-08-757-653-90
21	48	96.0	831	4	US-08-520-946-90
22	48	96.0	831	4	US-09-096-399-2
23	48	96.0	831	4	US-08-021-623C-6
24	48	96.0	831	4	US-09-096-399-4
25	48	96.0	831	5	PCT-US95-1418-5
26	48	96.0	831	5	PCT-US95-15327-5
27	48	96.0	831	5	PCT-US95-14418-4

28	45	90.0	597	5	PCT-US95-15327-4	Sequence 4, Appl1
29	45	90.0	810	4	US-09-587-856-2	Sequence 2, Appl1
30	45	90.0	810	4	US-09-777-537-2	Sequence 2, Appl1
31	45	90.0	810	4	US-09-777-538-2	Sequence 2, Appl1
32	45	90.0	832	1	US-07-977-434-2	Sequence 2, Appl1
33	45	90.0	832	1	US-08-156-020-2	Sequence 4, Appl1
34	45	90.0	832	1	US-08-156-020-6	Sequence 2, Appl1
35	45	90.0	832	1	US-08-156-020-6	Sequence 6, Appl1
36	45	90.0	832	1	US-08-156-020-10	Sequence 8, Appl1
37	45	90.0	832	1	US-08-073-384C-4	Sequence 10, Appl1
38	45	90.0	832	1	US-08-254-359A-4	Sequence 4, Appl1
39	45	90.0	832	1	US-08-483-043-4	Sequence 4, Appl1
40	45	90.0	832	1	US-08-458-819-2	Sequence 2, Appl1
41	45	90.0	832	1	US-08-481-238-4	Sequence 4, Appl1
42	45	90.0	832	2	US-08-471-066B-4	Sequence 4, Appl1
43	45	90.0	832	2	US-08-484-956-4	Sequence 4, Appl1
44	45	90.0	832	2	US-08-757-653-4	Sequence 4, Appl1
45	45	90.0	832	2	US-08-757-653-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-07-977-434-6
: Sequence 6, Application US/07977434
: Patent No. 546591
: GENERAL INFORMATION:
: APPLICANT: Gelband, David H.
: APPLICANT: Abramson, Richard D.
: TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
: TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
: NUMBER OF SEQUENCES: 38
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street
: CITY: Nutley
: STATE: New Jersey
: ZIP: 07110-1199
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Macintosh
: OPERATING SYSTEM: 7
: SOFTWARE: WordPerfect 2.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/977,434
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 590,490
: FILING DATE: 28-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 590,466
: FILING DATE: 28-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 590,213
: FILING DATE: 28-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 523,394
: FILING DATE: 15-MAY-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 143,441
: FILING DATE: 12-JAN-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 063,509
: FILING DATE: 17-JUN-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 899,241
: FILING DATE: 22-AUG-1986
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 746,121
: FILING DATE: 15-AUG-1991
: PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseri
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 5466591 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-977-434-6

Query Match 96.0%; Score 48; DB 1; Length 830;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSTPYEE 11
DB 676 LSQELSTPYEE 686

RESULT 2
US-08-458-819-6
Sequence 6, Application US/08458819
Patent No. 5795762
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: Wordperfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,819
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,434
FILING DATE: 23-FEB-1993
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseri
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 5795762 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-819-6

Query Match 96.0%; Score 48; DB 1; Length 830;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSTPYEE 11
DB 676 LSQELSTPYEE 686

RESULT 3
PCT-US91-07035-6
Sequence 6, Application PC/US9107035
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Celus Corporation
STREET: 1400 Fifty-third Street
CITY: Emeryville
STATE: California
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/07035
FILING DATE: 19910930
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Sias Ph.D., Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: Case No. 2580
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-420-3300
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCF-US91-07035-6

Query Match 96.0%; Score 48; DB 5; Length 830;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKSLIPYEE 11
Db 676 LSGELSIYEE 686

RESULT 4
US-08-073-384C-5
Sequence 5, Application US/08073384C
Patent No. 5541311
GENERAL INFORMATION:
APPLICANT: Dahlberg, James E.
APPLICANT: Lyamichev, Victor I.

APPLICANT: Brown, Mary Ann D.
TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
TITLE OF INVENTION: POLYMERASE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: HAVERSTOCK, MEDLEN & CARROLL,
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,384C
FILING DATE: 04-JUN-1993
CLASSIFICATION: 536
APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-00613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 841 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-073-384C-5

Query Match 96.0%; Score 48; DB 1; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKSLIPYEE 11
Db 677 LSGELSIYEE 687

RESULT 5
US-08-254-359A-5
Sequence 5, Application US/08254359A
Patent No. 5614402
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROWN, MARY ANN D.
TITLE OF INVENTION: 5' NUCLEASES DERIVED FROM THERMOSTABLE
TITLE OF INVENTION: DNA POLYMERASE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL,
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,359A

;; FILING DATE: 435
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/073,384
;; FILING DATE: 06-JUN-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/986,330
;; FILING DATE: 07-DEC-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CARROLL, PETER G.
;; REGISTRATION NUMBER: 32,837
;; REFERENCE/DOCKET NUMBER: FORS-01000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 831 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-254-359A-5

Query Match 96.0%; Score 48; DB 1; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
||| |||||

Db 677 LSGELSIPEE 687

RESULT 6
US-08-483-043-5
; Sequence 5, Application US/08483043
; Patent No. 5691142
; GENERAL INFORMATION:
; APPLICANT: Dahlberg, James E.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,043
; FILING DATE: 06-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,384
; FILING DATE: 04-JUN-1993
; APPLICATION NUMBER: US 07/986,330
; FILING DATE: 07-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-00613
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 5:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 831 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-483-043-5

Query Match 96.0%; Score 48; DB 1; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
||| |||||

Db 677 LSGELSIPEE 687

RESULT 7
US-08-481-238-5
; Sequence 5, Application US/08481238
; Patent No. 5795763
; GENERAL INFORMATION:
; APPLICANT: DAHLBERG, JAMES E.
; APPLICANT: LYAMICHEV, VICTOR I.
; APPLICANT: BROW, MARY ANN D.
; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
; TITLE OF INVENTION: POLYMERASE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,238
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: FORS-01798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 831 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-481-238-5

Query Match 96.0%; Score 48; DB 1; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
||| |||||

Db 677 LSGELSIPEE 687

RESULT 8
US-08-471-066B-5
; Sequence 5, Application US/08471066B
; Patent No. 5837450

GENERAL INFORMATION:
APPLICANT: Dahlberg, James E.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: 5' Nucleases Derived From Thermostable
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,066B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-01800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-471-066B-5

Query Match 96.0%; Score 48; DB 2; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSEXLSIPYEE 11
11 | | | | | | | | | |
Db 677 LSGELSLIPYEE 687

RESULT 9
US-08-484-956-5
Sequence 5, Application US/08484956
Patent No. 5843654
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: BROW, MARY ANN D.
APPLICANT: OLDENBURG, MARY C.
APPLICANT: HEISLER, LAURA
TITLE OF INVENTION: DETECTION OF P53 MUTATIONS
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVENSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO

STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, J. PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-956-5

Query Match 96.0%; Score 48; DB 2; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSEXLSIPYEE 11
11 | | | | | | | | | |
Db 677 LSGELSLIPYEE 687

RESULT 10
US-08-757-653-5
Sequence 5, Application US/08757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-757-653-5

Query Match 96.0%; Score 48; DB 2; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
Db 677 LSXELSIPEE 687

RESULT 11
US-08-599-491-5
Sequence 5, Application US/08599491
Patent No. 5846717
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: GROTELESCHEN HALL, JEFF S.
APPLICANT: LYAMICHEV, VICTOR
APPLICANT: OLIVE, DAVID M.
APPLICANT: PRUDENT, JAMES R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
TITLE OF INVENTION: INVADER-DIRECTED CLEAVAGE
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,491
FILING DATE: 23-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: P-40,027
REFERENCE/DOCKET NUMBER: FORS-01802
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-599-491-5

Query Match 96.0%; Score 48; DB 2; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
Db 677 LSXELSIPEE 687

RESULT 12
US-08-756-386-5
Sequence 5, Application US/08756386
Patent No. 5985557
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: LYAMICHEV, VICTOR I.
TITLE OF INVENTION: Invasive Cleavage of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-756-386-5

Query Match 96.0%; Score 48; DB 2; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
Db 677 LSXELSIPEE 687

RESULT 13
US-08-823-516-5
Sequence 5, Application US/08823516
Patent No. 5994069

```

APPLICANT: Grotelueschen Hall, Jeff S.
APPLICANT: Lyamichev, Victor
APPLICANT: Olive, David M.
APPLICANT: Prudent, James R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
TITLE OF INVENTION: INVADER-DIRECTED CLEAVAGE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/582,853A
FILING DATE: 12-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02306
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SPO ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: protein

US-08-682-853A-5

Query Match          96.0%; Score 48; DB 3; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Caps 0

CY      1 15X8LS1YEEV 11
Db      677 15GELSL1YEEV 687

RESULT 15
US-08-759-038-5
Sequence 5, Application US/08759038
Patent No. 6090543
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Brow, Mary Ann D.
APPLICANT: Dahlberg, James E.
TITLE OF INVENTION: Cleavage of Nucleic Acids
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,038
FILING DATE: 02-DEC-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02574
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO.: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-759-038-5

Query Match 96.0%; Score 48; DB 3; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.34;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 LSGELSTPYEE 11
Db 677 LSGELSTPYEE 687

Search completed: January 15, 2003, 12:38:23
Job time : 10.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: January 15, 2003, 11:20:34 ; Search time 10.8571 Seconds

(without alignments)
20.140 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSIPEE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCRUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	11	10 US-09-823-649A-4	Sequence 4, Appl1
2	48	96.0	11	10 US-09-823-649A-9	Sequence 9, Appl1
3	48	96.0	11	10 US-09-823-649A-12	Sequence 12, Appl1
4	48	96.0	11	10 US-09-823-649A-14	Sequence 14, Appl1
5	48	96.0	831	9 US-10-033-297-5	Sequence 5, Appl1
6	48	96.0	831	9 US-10-081-806-5	Sequence 5, Appl1
7	45	90.0	11	10 US-09-823-649A-3	Sequence 3, Appl1
8	45	90.0	11	10 US-09-823-649A-8	Sequence 8, Appl1
9	45	90.0	11	10 US-09-823-649A-10	Sequence 10, Appl1
10	45	90.0	11	10 US-09-823-649A-11	Sequence 11, Appl1
11	45	90.0	11	10 US-09-823-649A-13	Sequence 13, Appl1
12	45	90.0	832	9 US-10-033-297-4	Sequence 4, Appl1
13	45	90.0	832	9 US-09-972-834-2	Sequence 2, Appl1
14	45	90.0	832	9 US-10-071-505-1	Sequence 1, Appl1
15	45	90.0	832	9 US-10-081-806-4	Sequence 4, Appl1
16	45	90.0	833	9 US-10-033-297-8	Sequence 8, Appl1
17	45	90.0	833	9 US-10-033-297-66	Sequence 66, Appl1
18	45	90.0	833	9 US-10-033-297-69	Sequence 69, Appl1
19	45	90.0	833	9 US-10-033-297-71	Sequence 71, Appl1

20	45	90.0	833	9	US-10-081-806-8	Sequence 8, Appl1
21	45	90.0	834	9	US-10-033-297-6	Sequence 6, Appl1
22	45	90.0	834	9	US-10-081-806-6	Sequence 6, Appl1
23	45	90.0	836	10	US-09-777-430A-8	Sequence 8, Appl1
24	45	90.0	836	10	US-09-777-430A-11	Sequence 11, Appl1
25	45	90.0	842	10	US-09-777-430A-15	Sequence 15, Appl1
26	45	90.0	842	10	US-09-777-430A-20	Sequence 20, Appl1
27	45	90.0	842	10	US-09-777-430A-23	Sequence 23, Appl1
28	45	90.0	842	10	US-09-777-430A-26	Sequence 26, Appl1
29	44	88.0	11	9	US-09-823-649A-2	Sequence 2, Appl1
30	40	80.0	560	9	US-10-071-505-2	Sequence 2, Appl1
31	40	80.0	830	9	US-10-071-505-3	Sequence 3, Appl1
32	39	78.0	11	12	US-10-053-632-3	Sequence 3, Appl1
33	39	78.0	11	12	US-10-053-632-14	Sequence 14, Appl1
34	39	78.0	11	12	US-10-052-417-3	Sequence 3, Appl1
35	39	78.0	11	12	US-10-052-417-14	Sequence 14, Appl1
36	38	76.0	11	12	US-10-053-632-2	Sequence 2, Appl1
37	38	76.0	11	12	US-10-053-632-13	Sequence 13, Appl1
38	38	76.0	11	12	US-10-052-417-2	Sequence 2, Appl1
39	38	76.0	11	12	US-10-052-417-13	Sequence 13, Appl1
40	37	74.0	499	10	US-09-735-169A-2	Sequence 2, Appl1
41	37	74.0	499	10	US-09-735-171A-2	Sequence 2, Appl1
42	34	68.0	126	10	US-09-734-017A-58	Sequence 58, Appl1
43	33	66.0	194	10	US-09-925-302-585	Sequence 585, Appl1
44	33	66.0	375	12	US-10-062-254-373	Sequence 373, Appl1
45	33	66.0	415	10	US-09-826-212-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-823-649A-4
; Sequence 4, Application US/09823649A
; Patent No. US2002012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfsstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuichi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoenbrunner, Nancy
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYM
; FILE REFERENCE: RPA1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 69/129,336
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence motif
; NAME/KEY: VARIANT
; LOCATION: (3)-(3)
; OTHER INFORMATION: X is Q or G
US-09-823-649A-4

Query Match 96.0%; Score 48; DB 10; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSXELSIPEE 11
Db 1 LSXELSIPEE 11

RESULT 2
US-09-823-649A-9

```
; Sequence 9, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoedrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERAS
; FILE REFERENCE: RPA1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus flavus
US-09-823-649A-9
```

```
Query Match          96.0%; Score 48; DB 10; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.001;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 LSXELSIPEE 11
   |||||||
Db 1 LSQELSIPEE 11
```

```
RESULT 3
US-09-823-649A-12
; Sequence 12, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoedrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERAS
; FILE REFERENCE: RPA1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus sp. sps17
US-09-823-649A-12
```

```
Query Match          96.0%; Score 48; DB 10; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.001;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 LSXELSIPEE 11
   |||||||
Db 1 LSQELSIPEE 11
```

```
RESULT 4
US-09-823-649A-14
; Sequence 14, Application US/09823649A
; Patent No. US20020012970A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Smith, Edward
; APPLICANT: Elfstrom, Carita
; APPLICANT: Gelfand, David
; APPLICANT: Higuchi, Russell
; APPLICANT: Myers, Thomas
; APPLICANT: Schoedrunner, Nancy
; APPLICANT: Wang, Alice
; TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYM
; FILE REFERENCE: RPA1006
; CURRENT APPLICATION NUMBER: US/09/823,649A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/198,336
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Thermus filiformis
US-09-823-649A-14
```

```
Query Match          96.0%; Score 48; DB 10; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.001;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 LSXELSIPEE 11
   |||||||
Db 1 LSQELSIPEE 11
```

```
RESULT 5
US-10-033-297-5
; Sequence 5, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
;                               Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/033,297
; FILING DATE: 12-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/350,597
; FILING DATE: 09-Jul-1999
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
```

APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
S1 ANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-033-297-5

Query Match 96.0%; Score 48; DB 9; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISXELSTPYEE 11
DB 677 USXELSTPYEE 687

RESULT 6
US-10-081-806-5
Sequence 5, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION: --
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lymichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
City: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/081,806
FILING DATE: 22-Feb-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-081-806-5

Query Match 96.0%; Score 48; DB 9; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISXELSTPYEE 11
DB 677 USXELSTPYEE 687

RESULT 7
US-09-823-649A-3
Sequence 3, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carla
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYME
FILE REFERENCE: BPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: sequence motif
US-09-823-649A-3

Query Match 90.0%; Score 45; DB 10; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0039;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISXELSTPYEE 11
DB 1 USXELSTPYEE 11

RESULT 8
US-09-823-649A-8
Sequence 8, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carla
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYME
FILE REFERENCE: BPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 11

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/350,597
FILING DATE: 09-JUL-1999
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020187486A1 Relevant
TOPOLOGY: No. US20020187486A1 Relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-033-297-4
Query Match 90.0%; Score 45; DB 9; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.48;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSXELSPIYEE 11
|||:|||||
Db 678 LSQELAIYEE 688
RESULT 13
US-09-972-834-2
Sequence 2, Application US/09972834
Publication No. US20020192663A1
GENERAL INFORMATION:
APPLICANT: Loeb, Lawrence A.
Hood, Leroy
Suzuki, Motohshi
TITLE OF INVENTION: Thermostable Polymerases Having Altered
Fidelity and Methods of Identifying and Using Same
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL & FLORES LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/972,834
FILING DATE: 04-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/978,806
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UW 2873
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)535-9001
TELEFAX: (619)535-9049
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-972-834-2
Query Match 90.0%; Score 45; DB 9; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.48;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSXELSPIYEE 11
|||:|||||
Db 678 LSQELAIYEE 688
RESULT 14
US-10-071-505-1
Sequence 1, Application US/10071505
Publication No. US20020197211A1
GENERAL INFORMATION:
APPLICANT: Davis, Maria
APPLICANT: Nelson, John
APPLICANT: Kumar, Shiv
APPLICANT: Finn, Patrick J.
APPLICANT: Nampalli, Salyam
APPLICANT: Flick, Parke
TITLE OF INVENTION: TAQ DNA Polymerase Having an Amino Acid Substitution at
TITLE OF INVENTION: E681 and Homologs Thereof Exhibiting Improved Salt
Tolerance
FILE REFERENCE: PB9904
CURRENT APPLICATION NUMBER: US/10/071,505
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/7148,012
PRIOR FILING DATE: 1999-08-10
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn v.6.1
SEQ ID NO: 1
LENGTH: 832
TYPE: PRT
ORGANISM: Thermus aquaticus
US-10-071-505-1
Query Match 90.0%; Score 45; DB 9; Length 832;
Best Local Similarity 81.8%; Pred. No. 0.48;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSXELSPIYEE 11
|||:|||||
Db 678 LSQELAIYEE 688
RESULT 15
US-10-081-806-4
Sequence 4, Application US/10081806
Publication No. US20020197623A1
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP

Tue Jan 21 14:45:58 2003

us-09-823-649a-4.rapb

Page 6

```

? STREET: 220 Montgomery Street, Suite 2200
? CITY: San Francisco
? STATE: California
? COUNTRY: United States Of America
? ZIP: 94104
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/Ms-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/10/081,806
? FILING DATE: 22-Feb-2002
? CLASSIFICATION: <Unknown>
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/756,386
? FILING DATE: <Unknown>
? APPLICATION NUMBER: US 08/682,853
? FILING DATE: 12-JUL-1996
? APPLICATION NUMBER: US 08/599,491
? FILING DATE: 24-JAN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Ingolia, Diane E.
? REGISTRATION NUMBER: 40,027
? REFERENCE/DOCKET NUMBER: FORS-02564
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 832 amino acids
? TYPE: amino acid
? STRANDEDNESS:-single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? :
? SEQUENCE DESCRIPTION: SEQ ID NO: 4:
?
US-10-081-806-4
?
Query Match          90.0%; Score 45; DB 9; Length 832;
Best Local Similarity 81.8%; Pred.No. 0.48;
Matches      9; Conservative    1; Mismatches     1; Indels      0; Gaps       0;
OY           1 LSKKELSIPEYE 11
              |||||
Db            678 LSQELAIPYEE 688

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:15:24 / Search time 12.1429 Seconds

(without alignments)
87.086 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSIPEE 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	831	2 S26675	DNA-directed DNA p
2	45	90.0	831	2 JX0359	DNA-directed DNA p
3	45	90.0	832	2 A33530	DNA-directed DNA p
4	40	80.0	1451	2 B86286	P91.15 protein -
5	39	78.0	229	2 S46896	hypothetical prote
6	39	78.0	634	2 C70133	conserved hypothet
7	38	76.0	850	2 G70332	alpha-galactosidas
8	38	76.0	457	2 AE3130	alpha-galactosidas
9	38	76.0	474	2 D98157	genome polyploidin
10	37	74.0	1136	2 J00494	hypothetical coile
11	37	74.0	4717	2 T41581	E6 protein - human
12	36	72.0	149	1 W6WJ31	thiazole biosynthe
13	36	72.0	327	2 JC7337	protoprotein convert
14	36	72.0	753	1 KXMSCL	hypothetical prote
15	36	72.0	1890	2 T04556	glutathione S-tran
16	35	70.0	208	2 P87631	glutathione S-tran
17	35	70.0	238	2 S50809	hypothetical prote
18	35	70.0	381	2 T30452	hypothetical prote
19	35	70.0	738	2 T27289	protoprotein convert
20	35	70.0	752	1 KXRRC1	DNA-directed DNA p
21	35	70.0	921	2 A40597	hypothetical prote
22	35	70.0	2894	2 C64474	glutathione S-tran
23	34	68.0	184	2 S74356	hypothetical prote
24	34	68.0	225	2 A97735	hypothetical prote
25	34	68.0	233	2 AB1440	hypothetical prote
26	34	68.0	233	2 A11082	hypothetical prote
27	34	68.0	235	2 G90229	thiamin biosynthes
28	34	68.0	326	2 S25321	endo-1,4-beta-gluc
29	34	68.0	359	2 AC1580	

30	34	68.0	359	2 A11226	endo-1,4-beta-gluc
31	34	68.0	417	2 E75148	hypothetical prote
32	34	68.0	433	2 G64312	hypothetical prote
33	34	68.0	597	2 T40331	hypothetical prote
34	34	68.0	859	2 B64430	DNA-directed RNA p
35	34	68.0	1062	2 B26330	hypothetical prote
36	34	68.0	1234	2 B36186	I factor protein 2
37	33	66.0	118	2 S74769	hypothetical prote
38	33	66.0	132	2 G84529	probable poly(a) b
39	33	66.0	153	2 AC1431	hypothetical prote
40	33	66.0	153	2 A11804	hypothetical prote
41	33	66.0	162	1 DESPE	phosphoribosylamin
42	33	66.0	197	2 B64415	hypothetical prote
43	33	66.0	228	2 A53028	isopenicillin-diphos
44	33	66.0	291	2 B84373	ornithine cyclodaa
45	33	66.0	360	2 H81411	probable periplasm

ALIGNMENTS

RESULT 1

S26675
DNA-directed DNA polymerase (EC 2.7.7.7) I - Thermus aquaticus
C:Species: Thermus aquaticus
C>Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
C:Accession: S26675; S24929
R:Akhmetzjanov, A.A.; Vakhitov, V.A.
Nucleic Acids Res. 20, 5839, 1992
A:Title: Molecular cloning and nucleotide sequence of the DNA polymerase gene from Th
A:Reference number: S26675; MUID:93087201; PMID:1454544
A:Accession: S26675
A:Status: translation not shown

A:Molecule type: DNA
A:Residues: 1-831 <AKU>
A:Cross-references: EMBL:X66105; NID:g48165; PIDN:CAA46900.1; PID:g48166
A:Note: the source is designated as Thermus flavus
C:Genetics:
A:Gene: polI
C:Superfamily: DNA-directed DNA polymerase I
C:Keywords: DNA binding; nucleotidyltransferase

Query Match 96.0%; Score 48; DB 2; Length 831;
Best local similarity 90.9%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
DB 677 LSXELSIPEE 687

RESULT 2

JX0359
DNA-directed DNA polymerase (EC 2.7.7.7) - Thermus aquaticus
C:Species: Thermus aquaticus
C>Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 07-May-1999
C:Accession: JX0359
R:Rishino, Y.; Ueno, T.; Miyagi, M.; Demori, T.; Immamura, M.; Tsunawawa, S.; Kato, I.
J. Biochem. 116, 1019-1024, 1994
A:Title: Overproduction of Thermus aquaticus DNA polymerase and its structural analys
A:Reference number: JX0359; MUID:95204371; PMID:7896728
A:Accession: JX0359
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-831 <ISH>
C:Superfamily: DNA-directed DNA polymerase I
C:Keywords: nucleotidyltransferase

Query Match 90.0%; Score 45; DB 2; Length 831;
Best local similarity 81.8%; Pred. No. 0.75;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSXELSIPEE 11

DB 678 LSQELAIPIYEE 688

RESULT 3

DNA-directed DNA polymerase (EC 2.7.7.7) I - *Thermus aquaticus*
 C:Species: *Thermus aquaticus*
 C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 18-Jun-1999
 C:Accession: A33530
 R:Lawyer, F.C.; Stoffel, S.; Saiki, R.K.; Myambo, K.; Drummond, R.; Gelfand, D.H.
 J. Biol. Chem. 264, 6427-6437, 1989
 A:Title: Isolation, characterization, and expression in *Escherichia coli* of the DNA polymerase I
 A:Reference number: A33530; MUID:89197950; PMID:2649500
 A:Accession: A33530
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-832 <LAW>
 A:Cross-references: GB:J04639; NID:g155128; PIDN:AAA27507.1; PID:g155129; GB:M26480
 C:Superfamily: DNA-directed DNA polymerase I
 C:Keywords: DNA binding; nucleic acid transferase

Query Match 90.0%; Score 45; DB 2; Length 832;
 Best Local Similarity 81.8%; Pred. No. 0.75;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPIYEE 11
 DB 678 LSQELAIPIYEE 688

RESULT 4

B86286
 P91L15 protein - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C:Accession: B86286
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, T.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalla, R.; Moore, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1451 <STO>
 A:Cross-references: GB:AE005172; NID:g5103820; PIDN:AAD39650.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 80.0%; Score 40; DB 2; Length 1451;
 Best Local Similarity 63.6%; Pred. No. 15;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPIYEE 11
 DB 491 LSNEISVPIYDK 501

RESULT 5

S46696
 hypothetical protein YHR071w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein H8025.1
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 29-Oct-1999
 C:Accession: S46696
 R:Lafrelle, P.

submitted to the EMBL Data Library, May 1994
 A:Description: The sequence of *S. cerevisiae* cosmid 8025.
 A:Reference number: S46696
 A:Accession: S46696

A:Molecule type: DNA
 A:Residues: 1-229 <LAT>
 A:Cross-references: EMBL:U00061; NID:g487943; PIDN:AAB68375.1; PID:g487944; MIPS:YHR0
 C:Genetics:
 A:Gene: SC0:PCLS
 A:Cross-references: SC0:S000113; MIPS:YHR071w
 A:Map position: 8K

Query Match 78.0%; Score 39; DB 2; Length 229;
 Best Local Similarity 63.6%; Pred. No. 2.8;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPIYEE 11
 DB 174 LNYELAIPIYDE 184

RESULT 6

C70133
 conserved hypothetical protein B80267 - Lyme disease spirochete
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C:Accession: C70133
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, K.; Wh
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
 A:Reference number: A70100; MUID:98065943; PMID:9403685
 A:Accession: C70133
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-634 <KLE>
 A:Cross-references: GB:AE001137; GB:AE000783; NID:g2688160; PIDN:AAC66681.1; PID:g268
 A:Experimental source: strain B31

Query Match 78.0%; Score 39; DB 2; Length 634;
 Best Local Similarity 70.0%; Pred. No. 9;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SXMSIPIYDK 11
 DB 43 SLEISVPIYKE 52

RESULT 7

G70332
 conserved hypothetical protein ag_367 - *Aquifex aeolicus*
 C:Species: *Aquifex aeolicus*
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C:Accession: G70332
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: G70332
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-850 <AOF>
 A:Cross-references: GB:AE000686; NID:g2983038; PIDN:AAC06556.1; PID:g2983047; GB:AE00
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: ag_367

Query Match 78.0%; Score 39; DB 2; Length 850;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELSTPYEE 11
 DB 227 ISRMDVPEE 237

RESULT 8

AE3130

alpha-galactosidase mela [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 17-May-2002

C:Accession: AE3130

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, I.

E:Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AE3130

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-457 <KUR>

A:Cross-references: GB:AE008689; PIDN:AA145459.1; PID:q17743164; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: mela

A:Map position: linear chromosome

C:Superfamily: melibiose-specific alpha-galactosidase

Query Match

Best local similarity 76.0%; Score 38; DB 2; Length 457;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELSTPYEE 11

DB 183 LAHDLTPYEE 193

RESULT 9

D98157

alpha-galactosidase (melibiase) [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002

C:Accession: D98157

R:Goodier, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, H.; Goldman,

A.; Liu, F.; Molism, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 3323-3328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:11743194

A:Accession: D98157

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-474 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK88782.1; PID:q15158531; GSPDB:GN00170.

C:Genetics:

A:Gene: AGR_L436

A:Map position: linear chromosome

C:Superfamily: melibiose-specific alpha-galactosidase

Query Match

Best local similarity 76.0%; Score 38; DB 2; Length 474;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELSTPYEE 11

DB 200 LAHDLTPYEE 210

RESULT 10

J00494

genome polypeptide - Orithogalum mosaic virus (fragment)

N:Contains: coat protein; nuclear inclusion protein NI6; nuclear inclusion protein NI
 C:Species: Orithogalum mosaic virus, ORMV
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 17-Nov-2000
 C:Accession: J00494

R:Burger, J.

submitted to JIPID, May 1990

A:Reference number: J00494

A:Accession: J00494

A:Molecule type: genomic RNA

A:Residues: 1-1136 <BUR>

C:Superfamily: tobacco etch virus genome polypeptide

C:Keywords: polypeptide

E:1-365/Product: nuclear inclusion protein NI6 (fragment) #status predicted <NI6>

E:366-883/Product: nuclear inclusion protein NI6 #status predicted <NI6>

E:884-1136/Product: coat protein #status predicted <CVP>

Query Match

Best local similarity 74.0%; Score 37; DB 2; Length 1136;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 WLSIPYEE 11

DB 698 KLSIPYEE 705

RESULT 11

T41581

hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000

C:Accession: T41581

R:Murphy, L.; Harris, D.; Wood, V.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: 222002

A:Accession: T41581

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4717 <MUR>

A:Cross-references: EMBL:AL031546; PIDN:CAA20864.1; GSPDB:GN00068; SPDB:SPCC737.08

A:Experimental source: strain 972h(-)

C:Genetics:

A:Gene: SPDB:SPCC737.08

A:Map position: 3

C:Superfamily: Saccharomyces probable membrane protein YKR106c

Query Match

Best local similarity 74.0%; Score 37; DB 2; Length 4717;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSXELSTPYEE 11

DB 2663 LTRKSLPYEE 2673

RESULT 12

W6WL31

E6 protein - human papillomavirus type 31

C:Species: human papillomavirus type 31

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999

C:Accession: A32444

R:Goldborough, M.D.; Disilvestre, D.; Temple, G.F.; Lorincz, A.T.

Virology 171, 306-311, 1989

A:Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-as

A:Reference number: A94398; MID:8929478; PMID:2545036

A:Accession: A32444

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-149 <GOL>

A:Cross-references: GB:J04353; NID:9333048; PIDN:AAA46950.1; PID:q459916

C:Comment: This protein may be involved in the oncogenic potential of this virus.

C:Keywords: DNA binding; early protein; zinc finger

F:30-66/Region: zinc finger CCCC motif
F:103-139/Region: zinc finger CCCC motif

Query Match 72.0% Score 36; DB 1; Length 149;
Best Local Similarity 63.6%; Pred. No. 6.9;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ISXSLIPYER 11
11 | 11111
DB 15 LSSALPIPYDE 25

RESULT 13

JC7337

thiazole biosynt.atic enzyme - Aspergillus oryzae
N:Alternate names: pyridoxamine resistance protein

C:Species: Aspergillus oryzae

C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 03-Aug-2001

C:Accession: JC7337

R:Kubodera, T.; Yamashita, N.; Nishimura, A.

Biosci. Biotechnol. Biochem. 64, 1416-1421, 2000

A:Title: Pyridoxamine resistance gene (ptrA) of Aspergillus oryzae: Cloning, characterization

A:Reference number: JC7337; MUID:20399355; PMID:10945258

A:Accession: JC7337

A:Molecule type: DNA

A:Residues: 1-327 <KUB>

A:Cross-references: GB:AF217503; PID:96694866; PIDN:AAF25444.1; PID:96694867

A:Experimental source: strain wild type

C:Genetics:

A:Gene: ptrA

A:Introns: 113/1

C:Superfamily: thiamin biosynthesis protein thl1

Query Match 72.0% Score 36; DB 2; Length 327;
Best Local Similarity 75.0%; Pred. No. 17;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELSIPYER 11
11 | 11111
DB 138 ELGVPEYER 145

RESULT 14

KXMSCI

proprotein convertase 1 (EC 3.4.21.93) precursor - mouse

N:Alternate names: furin homolog PC1; kexin homolog PC1; prohormone cleavage enzyme; pro

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-May-2000

C:Accession: JX0171; S19165; A39604; A35571; A39002; A37951; A46622

R:Nakayama, K.; Hosaka, M.; Hattuzawa, K.; Murakami, K.

J. Biochem. 109, 803-806, 1991

A:Title: Cloning and functional expression of a novel endoprotease involved in prohormone

A:Reference number: JX0171; MUID:92041727; PMID:1657897

A:Accession: JX0171

A:Molecule type: mRNA

A:Residues: 1-753 <NAK>

A:Experimental source: pituitary AtT-20 cells

R:Nakayama, K.

submitted to the EMBL Data Library December 1990

A:Reference number: S19165

A:Accession: S19165

A:Molecule type: mRNA

A:Residues: 1-111, 'P', 113-116, 'P', 118-121, 'T', 123-127, 'H', 129-281, 'K', 283-731, 'E', 733-75

A:Cross-references: EMBL:X57088; NID:950054; PIDN:CAA40368.1; PID:950055

R:Korner, J.; Chun, J.; Harter, D.; Axel, R.

Proc. Natl. Acad. Sci. U.S.A. 88, 6834-6838, 1991

A:Title: Isolation and functional expression of a mammalian prohormone processing enzyme

A:Reference number: A39604; MUID:91319778; PMID:1862107

A:Accession: A39604

A:Molecule type: mRNA

A:Residues: 1-609 <KOR>

A:Cross-references: GB:M69196

R:Seidah, N.G.; Gaspar, L.; Mion, P.; Marcinkiewicz, M.; Mbikay, M.; Chretien, M.

DNA Cell Biol. 9, 415-424, 1990

A:Title: cDNA sequence of two distinct pituitary proteins homologous to kex2 and fur1

A:Reference number: A35571; MUID:91000356; PMID:2169760

A:Accession: A35571

A:Molecule type: mRNA

A:Residues: 214-478 <SEL>

A:Note: The authors gave the codon for residue 330-Ser as CTC

R:Smekens, S.P.; Avrich, A.S.; Lamendola, J.; Chan, S.J.; Steiner, D.F.

Proc. Natl. Acad. Sci. U.S.A. 88, 340-344, 1991

A:Title: Identification of a cDNA encoding a second putative prohormone convertase re

A:Reference number: A39002; MUID:91110525; PMID:1988934

A:Accession: A39002

A:Molecule type: mRNA

A:Residues: 1-22, 'S', 24-753 <SMR>

A:Cross-references: GB:M58507; NID:9200242; PIDN:AAA39896.1; PID:9200243

A:Experimental source: pituitary AtT-20 cells

R:Seidah, N.G.; Marcinkiewicz, M.; Benjannet, S.; Gaspar, L.; Beaubien, G.; Mallet, N

Mol. Endocrinol. 5, 111-122, 1991

A:Title: Cloning and primary sequence of a mouse candidate prohormone convertase PC1

A:Reference number: A37951; MUID:91203919; PMID:2017186

A:Accession: A37951

A:Molecule type: mRNA

A:Residues: 1-360, 'S', 362-363, 'P', 365-753 <SE2>

A:Cross-references: GB:M58589

R:Zhou, Y.; Lindberg, I.

J. Biol. Chem. 268, 5615-5623, 1993

A:Title: Purification and characterization of the prohormone convertase PC1(PC3).

A:Reference number: A46622; MUID:93194858; PMID:8449925

A:Accession: A46622

A:Molecule type: Protein

A:Residues: 111-120 <ZHO>

C:Comment: This protein is a member of a family of subtilisin-like proteinases respon

C:Genetics:

A:Map position: 13C

C:Superfamily: kexin; subtilisin homology

C:Keywords: glycoprotein; hydrolase; serine proteinase

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-110/Domain: propeptide #status predicted <PRO>

F:111-753/Product: prohormone-processing proteinase PC1 #status predicted <MAT>

F:158-396/Domain: subtilisin homology <SHR>

F:167-208,382/Active site: Asp, His, Ser #status predicted

F:173-401,645/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 72.0% Score 36; DB 1; Length 753;
Best Local Similarity 77.8%; Pred. No. 44;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SXMSIPYE 10
1 | 1111111
DB 690 SAKSLIPYE 698

RESULT 15

T04556

hypothetical protein F28J12.260 - Arabidopsis thaliana

N:Alternate names: hypothetical protein F28J12.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T04556; T04852

R:Bevan, M.; Hilbert, H.; Brann, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancro

submitted to the Protein Sequence Database, February 1998

A:Reference number: Z15377

A:Accession: T04556

A:Molecule type: DNA

A:Residues: 1-1890 <BEV>

A:Cross-references: EMBL:AL021710

A:Experimental source: cultivar Columbia; BAC clone F28J12

R:Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; May

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15387

A:Accession: T04852

A: Molecule type: DNA
A: Residues: 1-1560 <BBW>
A: Cross-references: EMBL:AL035526
A: Experimental source: cultivar Columbia; BAC clone F28A21
C: Genetics:
A: Map position: 4
A: Introns: 54/2; 109/1; 150/2; 194/3; 324/3; 1470/1; 1785/1
A: Note: F28J12.260; F28A21.10

Query Match 72.0%; Score 36; DB 2; Length 1890;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SXELSPYEE 11
| | | | |
Db 1137 SLEESVPYEE 1146

Search completed: January 15, 2003, 12:37:19
Job time: 14.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 15, 2003, 12:36:21 : Search time 26 Seconds

(without alignments)
129.748 Million cell updates/sec

Title: US-09-823-649A-3
Perfect score: 54
Sequence: 1 LSOEHAIPYEE 11

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters: **
-MODE=frame+ p2n.model -DEV=xlh
-O=/cg2n2.1/USPTO.spool/US09823649/ynat.14012003.151002.29127/app.query.fasta.1.1393
-DB-ISSUED.Patents.NA -QWMT=fastap -SUFFIX=ini -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649 @CGN1.1.35 @rnat.14012003.151002.29127 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -IARGUMENTY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMECUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued.Patents.NA:*

1: /cg2n2.6/ptodata/1/ina/5A.COMB.seq:*
2: /cg2n2.6/ptodata/1/ina/5A.COMB.seq:*
3: /cg2n2.6/ptodata/1/ina/5A.COMB.seq:*
4: /cg2n2.6/ptodata/1/ina/5A.COMB.seq:*
5: /cg2n2.6/ptodata/1/ina/5A.COMB.seq:*
6: /cg2n2.6/ptodata/1/ina/5A.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	1600	1 US-08-073-384C-12	Sequence 12, Appl
2	54	100.0	1600	1 US-08-254-359A-12	Sequence 12, Appl
3	54	100.0	1600	1 US-08-483-043-12	Sequence 12, Appl
4	54	100.0	1600	1 US-08-481-238-12	Sequence 12, Appl
5	54	100.0	1600	2 US-08-471-066B-12	Sequence 12, Appl
6	54	100.0	1600	2 US-08-484-956-12	Sequence 12, Appl
7	54	100.0	1600	2 US-08-757-653-12	Sequence 12, Appl
8	54	100.0	1600	2 US-08-599-491-12	Sequence 12, Appl
9	54	100.0	1600	2 US-08-756-386-12	Sequence 12, Appl
10	54	100.0	1600	2 US-08-823-516-12	Sequence 12, Appl
11	54	100.0	1600	3 US-08-682-853A-12	Sequence 12, Appl
12	54	100.0	1600	3 US-08-759-038-12	Sequence 12, Appl

13	54	100.0	1600	3 US-08-758-314-12	Sequence 12, Appl
14	54	100.0	1600	4 US-09-350-309-12	Sequence 12, Appl
15	54	100.0	1600	4 US-08-520-946-12	Sequence 12, Appl
16	54	100.0	1682	2 US-09-096-399-1	Sequence 1, Appl
17	54	100.0	1686	2 US-08-648-657-1	Sequence 1, Appl
18	54	100.0	1686	2 US-08-648-657-1	Sequence 1, Appl
19	54	100.0	1686	2 US-08-648-657-2	Sequence 14, Appl
20	54	100.0	1696	2 US-09-096-399-3	Sequence 3, Appl
21	54	100.0	1794	5 PCT-US95-14418-3	Sequence 3, Appl
22	54	100.0	2499	1 PCT-US95-15327-3	Sequence 3, Appl
23	54	100.0	2499	1 US-07-977-434-1	Sequence 1, Appl
24	54	100.0	2499	1 US-08-458-819-1	Sequence 1, Appl
25	54	100.0	2499	3 US-07-602-848B-1	Sequence 1, Appl
26	54	100.0	2499	4 US-09-587-856-1	Sequence 1, Appl
27	54	100.0	2499	4 US-09-777-537-1	Sequence 1, Appl
28	54	100.0	2499	4 US-09-777-538-1	Sequence 1, Appl
29	54	100.0	2499	5 PCT-US91-07035-1	Sequence 1, Appl
30	54	100.0	2502	1 US-08-073-384C-7	Sequence 7, Appl
31	54	100.0	2502	1 US-08-073-384C-21	Sequence 21, Appl
32	54	100.0	2502	1 US-08-254-359A-7	Sequence 7, Appl
33	54	100.0	2502	1 US-08-254-359A-21	Sequence 21, Appl
34	54	100.0	2502	1 US-08-483-043-7	Sequence 7, Appl
35	54	100.0	2502	1 US-08-483-043-21	Sequence 21, Appl
36	54	100.0	2502	1 US-08-481-238-7	Sequence 7, Appl
37	54	100.0	2502	1 US-08-481-238-21	Sequence 21, Appl
38	54	100.0	2502	2 US-08-471-066B-7	Sequence 7, Appl
39	54	100.0	2502	2 US-08-471-066B-21	Sequence 21, Appl
40	54	100.0	2502	2 US-08-484-956-7	Sequence 7, Appl
41	54	100.0	2502	2 US-08-484-956-21	Sequence 21, Appl
42	54	100.0	2502	2 US-08-757-653-7	Sequence 7, Appl
43	54	100.0	2502	2 US-08-757-653-21	Sequence 21, Appl
44	54	100.0	2502	2 US-08-599-491-7	Sequence 7, Appl
45	54	100.0	2502	2 US-08-599-491-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-073-384C-12
Sequence 12, Application US/08073384C
Patent No. 5541311
GENERAL INFORMATION:
APPLICANT: Dahlberg, James E.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
TITLE OF INVENTION: POLYMERASE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSOCK, MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073.384C
FILING DATE: 04-JUN-1993
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-00613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410

TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-073-384C-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-3 (1-11) x US-08-073-384C-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11

Db 1133 CTCGCCAGAGACTAGCATCCCTTACGAGAG 1165

RESULT 2
US-08-254-359A-12

; Sequence 12, Application US/08254359A

; Patent No. 5614402

; GENERAL INFORMATION:

; APPLICANT: DAHLBERG, JAMES E.

; APPLICANT: LYAMICHEV, VICTOR I.

; APPLICANT: BROW, MARY ANN D.

; TITLE OF INVENTION: 5,2 NUCLEASES DERIVED FROM THERMOSTABLE

; TITLE OF INVENTION: DNA POLYMERASE

; NUMBER OF SEQUENCES: 40

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: HAVERTOCK, MEDLEN & CARROLL

; STREET: 220 MONTGOMERY STREET, SUITE 2200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/254,359A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/073,384

; FILING DATE: 06-JUN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/986,330

; FILING DATE: 07-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: CARROLL, PETER G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: FORS-01000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1600 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-254-359A-12

Alignment Scores:

Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-3 (1-11) x US-08-254-359A-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11

Db 1133 CTCGCCAGAGACTAGCATCCCTTACGAGAG 1165

RESULT 3

US-08-483-043-12

; Sequence 12, Application US/08483043

; Patent No. 5691142

; GENERAL INFORMATION:

; APPLICANT: Dahlberg, James E.

; APPLICANT: Lyamichev, Victor I.

; APPLICANT: Brow, Mary Ann D.

; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA

; TITLE OF INVENTION: POLYMERASE

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: HAVERTOCK, MEDLEN & CARROLL

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,043

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/073,384

; FILING DATE: 04-JUN-1993

; APPLICATION NUMBER: US 07/986,330

; FILING DATE: 07-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Carroll, Peter G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: FORS-00613

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/705-8410

; TELEFAX: 415/397-8338

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1600 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-483-043-12

Alignment Scores:

Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-3 (1-11) x US-08-483-043-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11

DB 1133 CTCCTCCAGAGACTAGCATCCCTTACGAGAG 1165

RESULT 4

US-08-481-238-12
Sequence 12, Application US/08481238
Patent No. 5795763

GENERAL INFORMATION:

APPLICANT: DAHLBERG, JAMES E.
APPLICANT: LYAMICHEV, VICTOR I.

TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
TITLE OF INVENTION: POLYMERASE
NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESSES:

ADDRESSER: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,238

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CARROLL, PETER G.

REGISTRATION NUMBER: 92,837

REFERENCE/DOCKET NUMBER: FONS-01798

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1600 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-481-238-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Caps: 0

US-09-823-649A-3 (1-11) x US-08-481-238-12 (1-1600)

QY 1 LeuserGlnGluAlaAlaIleProTyrGluGlu 11

DB 1133 CTCCTCCAGAGACTAGCATCCCTTACGAGAG 1165

RESULT 5

US-08-471-066B-12
Sequence 12, Application US/08471066B
Patent No. 5837450

GENERAL INFORMATION:

APPLICANT: DAHLBERG, JAMES E.

APPLICANT: LYAMICHEV, VICTOR I.

APPLICANT: BROW, MARY ANN D.

TITLE OF INVENTION: 5' Nucleases Derived From Thermostable

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESSES:

ADDRESSER: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,066B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/254,359

FILING DATE: 06-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/073,384

FILING DATE: 04-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/986,330

FILING DATE: 07-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FONS-01800

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1600 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-471-066B-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Caps: 0

US-09-823-649A-3 (1-11) x US-08-471-066B-12 (1-1600)

QY 1 LeuserGlnGluAlaAlaIleProTyrGluGlu 11

DB 1133 CTCCTCCAGAGACTAGCATCCCTTACGAGAG 1165

US-08-484-956-12
Sequence 12, Application US/08484956
Patent No. 5843654

GENERAL INFORMATION:

APPLICANT: DAHLBERG, JAMES E.

APPLICANT: LYAMICHEV, VICTOR I.

APPLICANT: BROW, MARY ANN D.

APPLICANT: OLSENBERG, MARY C.

TITLE OF INVENTION: DETECTION OF p53 MUTATIONS

NUMBER OF SEQUENCES: 114

CORRESPONDENCE ADDRESSES:

ADDRESSER: HAVERSTOCK, MEDLEN & CARROLL

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, J., PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-956-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-823-649A-3 (1-11) x US-08-484-956-12 (1-1600)
QY 1 LeuSerGInGluLeuAlaIleProTyGInGlu 11
Db 1133 CTCCTCCAGAGCTAGCCATCCCTTACGAGAG 1165

RESULT 7
US-08-757-653-12
Sequence 12, Application US/08/757653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-653-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-823-649A-3 (1-11) x US-08-757-653-12 (1-1600)
QY 1 LeuSerGInGluLeuAlaIleProTyGInGlu 11
Db 1133 CTCCTCCAGAGCTAGCCATCCCTTACGAGAG 1165

RESULT 8
US-08-599-491-12
Sequence 12, Application US/08/599491
Patent No. 5846717
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: GROTEJUESCHEN HALL, JEFF S.
APPLICANT: LYAMICHEV, VICTOR
APPLICANT: OLIVE, DAVID M.
APPLICANT: PADDENT, JAMES R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,491
FILING DATE: 23-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: P-40,027
REFERENCE/DOCKET NUMBER: FORS-01802
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-599-491-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-823-649a-3 (1-11) x US-08-599-491-12 (1-1600)
QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
Db 1133 CTCCTCCAGAGCTACGATCCCTTACGAGAG 1165

RESULT 9
US-08-756-386-12
Sequence 12, Application US/08756386
Patent No. 5985557

GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichiev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-756-386-12

Alignment Scores:
Pred. No.: 0.0143 Length: 1600

Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-823-649a-3 (1-11) x US-08-756-386-12 (1-1600)
QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
Db 1133 CTCCTCCAGAGCTACGATCCCTTACGAGAG 1165

RESULT 10
US-08-823-516-12
Sequence 12, Application US/08823516
Patent No. 5994069

GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichiev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brown, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/756,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-823-516-12

Alignment Scores:

Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-823-649a-3 (1-11) x US-08-823-516-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGlnGlu 11

Db 1133 CTCCTCCAGAGCTAGCCATCCCTTACGAGAG 1165

RESULT 11

US-08-682-853a-12
Sequence 12, Application US/08682853A
Patent No. 6001567

GENERAL INFORMATION:

APPLICANT: Brown, Mary Ann D.
APPLICANT: Groteschusen Hall, Jeff S.
APPLICANT: Lyamichov, Victor I.
APPLICANT: Olive, David M.
APPLICANT: Prudent, James R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
TITLE OF INVENTION: INVADER-DIRECTED CLEAVAGE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/682,853A
FILING DATE: 12-JUL-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
TELECOMMUNICATION INFORMATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02306

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
US-08-682-853a-12

Alignment Scores:

Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-823-649a-3 (1-11) x US-08-682-853a-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGlnGlu 11

Db 1133 CTCCTCCAGAGCTAGCCATCCCTTACGAGAG 1165

RESULT 12

US-08-759-038-12
Sequence 12, Application US/08759038
Patent No. 6090543

GENERAL INFORMATION:

APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichov, Victor I.
APPLICANT: Brown, Mary Ann D.
APPLICANT: Dahlberg, James E.
TITLE OF INVENTION: Cleavage of Nucleic Acids
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,038
FILING DATE: 02-DEC-1996
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: FORS-02574
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
US-08-759-038-12

Alignment Scores:

Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-823-649a-3 (1-11) x US-08-759-038-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGlnGlu 11

Db 1133 CTCCTCCAGAGCTAGCCATCCCTTACGAGAG 1165

RESULT 13

US-08-758-314-12

Sequence 12, Application US/08758314
Patent No. 6090606
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natascha
TITLE OF INVENTION: Improved Cleavage Agents
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,314
FILING DATE: 02-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/662,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02575
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-758-314-12
Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-09-823-649A-3 (1-11) x US-08-758-314-12 (1-1600)
QY 1 LeuserGIngluLeuAlaIleProTyrGluGln 11
|||||
Db 1133 CTCTCCAGAGACTAGCAGCATCTTACGAGAG 1165
RESULT 14
US-09-350-309-12
Sequence 12, Application US/09350309
Patent No. 6348314
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
Hall, Jeff G.
Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage of Nucleic Acids
NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/350,309
FILING DATE: 09-JUL-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/756,386
FILING DATE: <unknown>
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-350-309-12
Alignment Scores:
Pred. No.: 0.0143 Length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-823-649A-3 (1-11) x US-09-350-309-12 (1-1600)
QY 1 LeuserGIngluLeuAlaIleProTyrGluGln 11
|||||
Db 1133 CTCTCCAGAGACTAGCAGCATCTTACGAGAG 1165
RESULT 15
US-08-520-946-12
Sequence 12, Application US/08520946
Patent No. 6572424
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: LYAMICHEV, VICTOR I.
APPLICANT: OLIVE, DAVID M.
TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/520,946
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01756
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-520-946-12

Alignment Scores:
Pred. No.: 0.0143 length: 1600
Score: 54.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 4

US-09-823-649a-3 (1-11) x US-08-520-946-12 (1-1600)

QY 1 LeuSerGlnGluLeuAlaIleProTyrGluGlu 11
|||||
Db 1133 CTCCTCCAGAGAGTAGCCATCCCTTAGAGAGAG 1165

Search completed: January 15, 2003, 12:58:42
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: January 15, 2003, 11:15:24 / Search time 12.1429 Seconds
(without alignments)
87.086 Million cell updates/sec

Title: US-09-823-649A-3
Perfect score: 54
Sequence: 1 LSQELAIPEE 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	54	100.0	831	2 JX0359	DNA-directed DNA p
2	54	100.0	832	2 A33530	DNA-directed DNA p
3	44	81.5	831	2 S26675	DNA-directed DNA p
4	42	77.8	229	2 S46696	hypothetical prote
5	39	72.2	850	2 G70332	conserved hypothet
6	38	70.4	326	2 S25321	thiamin biosynthes
7	38	70.4	1451	2 B86286	F9L1.13 protein -
8	37	68.5	417	2 B59096	hypothetical prote
9	37	68.5	452	2 B59096	hypothetical prote
10	37	68.5	457	2 AE3130	alpha-galactosidas
11	37	68.5	474	2 D98157	alpha-galactosidas
12	36	66.7	327	2 JC7337	thiazole biosynthes
13	36	66.7	412	2 T34415	hypothetical prote
14	36	66.7	921	2 A40597	DNA-directed DNA p
15	35	64.8	59	2 C86902	hypothetical prote
16	35	64.8	149	1 W6W131	B6 protein - human
17	35	64.8	208	2 F87631	glutathione S-tran
18	35	64.8	240	2 G89782	glutathione S-tran
19	35	64.8	402	2 G72516	probable acetyl-CO
20	35	64.8	634	2 C70133	conserved hypothet
21	35	64.8	1448	2 T08526	DNA primase tracr2
22	35	64.8	1448	2 S37669	tracr-2 protein - E
23	35	64.8	1469	2 H96622	probable ABC trans
24	34	63.0	220	2 C87601	glutathione S-tran
25	34	63.0	222	2 G70440	cytidylate kinase
26	34	63.0	233	2 AB1440	hypothetical prote
27	34	63.0	233	2 A11082	hypothetical prote
28	34	63.0	238	2 S50809	ribosomal protein
29	34	63.0	273	2 C83447	conserved hypothet

30	34	63.0	285	2 S29368	translation initia
31	34	63.0	309	2 C83754	hypothetical prote
32	34	63.0	342	2 AF2177	glucokinase (impor
33	34	63.0	382	2 S01339	exo-alpha-sialidas
34	34	63.0	382	2 S32148	exo-alpha-sialidas
35	34	63.0	385	2 C75221	hypothetical prote
36	34	63.0	390	2 C75221	molycoprotein bios
37	34	63.0	415	2 B84397	Nifs protein, clas
38	34	63.0	458	2 G72377	aspartate ammonia-
39	34	63.0	458	2 G63317	conserved hypothet
40	34	63.0	629	2 C86149	T186.18 protein -
41	34	63.0	1136	2 JQ0494	genome polypotein
42	34	63.0	1324	2 S06187	RNA2 polypotein -
43	33	61.1	105	2 D72452	hypothetical prote
44	33	61.1	153	2 AC1431	hypothetical prote
45	33	61.1	153	2 A11804	hypothetical prote

ALIGNMENTS

RESULT 1
JX0359
DNA-directed DNA polymerase (EC 2.7.7.7) - Thermus aquaticus
C:Species: Thermus aquaticus
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 07-May-1999
C:Accession: JX0359
R:Rhino, Y.; Ueno, T.; Miyagi, M.; Uemori, T.; Imamura, M.; Tsunawaka, S.; Kato, I.
J. Biochem. 116, 1019-1024, 1994
A:Title: Overproduction of Thermus aquaticus DNA polymerase and its structural analysis
A:Reference number: JX0359; M01D:95204371; PMID:7896728
A:Accession: JX0359
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-831 <ISH>
C:Superfamily: DNA-directed DNA polymerase I
C:Keywords: nucleotidyltransferase

Query Match 100.0%; Score 54; DB 2; Length 831;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
|||||

DB 678 LSQELAIPEE 688

RESULT 2
A33530
DNA-directed DNA polymerase (EC 2.7.7.7) I - Thermus aquaticus
C:Species: Thermus aquaticus
C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 18-Jun-1999
C:Accession: A33530
R:Liwyer, F.C.; Stoffel, S.; Saiki, R.K.; Myambo, K.; Drummond, R.; Gelfand, D.H.
J. Biol. Chem. 264, 6427-6437, 1989
A:Title: Isolation, characterization, and expression in Escherichia coli of the DNA p

A:Reference number: A33530; M01D:89197950; PMID:2649500
A:Accession: A33530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-832 <LAW>
A:Cross-references: GB:J04639; NID:9155128; PIDN:AA27507.1; PID:9155129; GB:M26480
C:Superfamily: DNA-directed DNA polymerase I
C:Keywords: DNA binding; nucleotidyltransferase

Query Match 100.0%; Score 54; DB 2; Length 832;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
|||||

DB 678 LSQELAIPEE 688

```

RESULT 3
S26675
DNA-directed DNA polymerase (EC 2.7.7.7) I - Thermus aquaticus
C:Species: Thermus aquaticus
C:Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
C:Accession: S26675; S24929
R:Khmetyanov, A.A.; Vakhitov, V.A.
Nucleic Acids Res. 20, 5839, 1992
A:Title: Molecular cloning and nucleotide sequence of the DNA polymerase gene from Therm
A:Reference number: S26675; MUID:93087201; PMID:1454544
A:Accession: S26675
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-831 <AKH>
A:Cross-references: EMBL:X66105; NID:948165; PIDN:CAA46900.1; PID:948166
A:Note: the source is designated as Thermus flavus
C:Genetics:
A:Gene: polI
C:Superfamily: DNA-directed DNA polymerase 1
C:Keywords: DNA binding; nucleotidyltransferase

Query Match      81.5%; Score 44; DB 2; Length 831;
Best local similarity 81.8%; Pred. No. 2.7;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
|||:|||||
DB 677 LSGELIPEE 687

RESULT 4
S46696
hypothetical protein YHR071w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein H8025.1
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 29-Oct-1999
C:Accession: S46696
R:Latreille, P.
Submitted to the EMBL Data Library, May 1994
A:Description: The sequence of S. cerevisiae cosmid 8025.
A:Reference number: S46696
A:Accession: S46696
A:Molecule type: DNA
A:Residues: 1-229 <LAT>
A:Cross-references: EMBL:U00061; NID:9487943; PIDN:AAB68375.1; PID:9487944; MIPS:YHR071w
C:Genetics:
A:Gene: SGD:PCLS
A:Cross-references: SGD:S000113; MIPS:YHR071w
A:Map position: 8R

Query Match      77.8%; Score 42; DB 2; Length 229;
Best local similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ISQELAIPEE 11
|:|||||:|
DB 174 LNELAIPEDE 184

RESULT 5
G70332
conserved hypothetical protein aq_367 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: G70332
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:9619666; PMID:9537320
C:Accession: G70332
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

```

```

A:Molecule type: DNA
A:Residues: 1-850 <AQ>
A:Cross-references: GB:AE000686; NID:92983038; PIDN:AAC06656.1; PID:92983047; GB:AE00
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_367

Query Match      72.2%; Score 39; DB 2; Length 850;
Best local similarity 54.5%; Pred. No. 26;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
|||:|||||
DB 227 ISQELAIPEE 237

RESULT 6
S25321
thiamin biosynthesis protein thl4 - yeast (Saccharomyces cerevisiae)
N:Alternate names: MOLI protein; protein G6620; protein TOR144w
C:Species: Saccharomyces cerevisiae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Jun-2000
C:Accession: S25321; S60435; S64453; S17019
R:Praekelt, U.M.; Meacock, P.A.
Yeast 8, 699-710, 1992
A:Title: MOLI, a Saccharomyces cerevisiae gene that is highly expressed in early stat
A:Reference number: S25321; MUID:93070608; PMID:1441749
A:Accession: S25321
A:Molecule type: DNA
A:Residues: 1-326 <PRA>
A:Cross-references: EMBL:X61669; NID:94588; PIDN:CAA43843.1; PID:94589
R:Skala, J.; Nawrocki, A.; Goffeau, A.
Yeast 11, 1421-1427, 1995
A:Title: The sequence of a 27 kb segment on the right arm of chromosome VII from Sacc

A:Reference number: S60435; MUID:96158062; PMID:8585325
A:Accession: S60435
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 205-326 <SKA>
A:Cross-references: EMBL:X65807; NID:91045249; PIDN:CAA59802.1; PID:91045250
A:Note: The nucleotide sequence was submitted to the EMBL data library, March 1995
R:Van Dyck, L.; Skala, J.; de Wergifosse, P.; Punnele, B.; Talla, E.; Nawrocki, A.;
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64428
A:Accession: S64428
A:Molecule type: DNA
A:Residues: 1-326 <VAN>
A:Cross-references: EMBL:X72929; NID:91323241; PIDN:CAA97157.1; PID:91323242; MIPS:YC
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:THI4; MOLI: ESP35
A:Cross-references: SGD:S0003376; MIPS:YGR144w
A:Map position: 7R
C:Superfamily: thiamin biosynthesis protein thl1
C:Keywords: transmembrane protein
F:68-84/Domain: transmembrane #status predicted <TM>

Query Match      70.4%; Score 38; DB 2; Length 326;
Best local similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 QELAIPEE 11
|||:|||||
DB 127 QELIPEED 135

RESULT 7
B86286
F9L1.15 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: B86286

```

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, A.; Authors: Salzborg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; PMID:21016719; PMID:11130712
 A:Accession: B86286
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1451 <STO>
 A:Cross-references: GB:AE005172; NID:q5103820; PIDN:AAD39650.1; CSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 70.4%; Score 38; DB 2; Length 1451;
 Best Local Similarity 54.5%; Pred. No. 73;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LSOELAPIYEE 11
 |||||:
 Db 491 LSNELSVPYDX 501

RESULT 8
 E75148
 Hypothetical protein PAB2116 - *Pyrococcus abyssi* (strain Orsay)
 C:Species: *Pyrococcus abyssi*
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: E75148
 R:Anonymous, Genoscope
 Submitted to the EMBL Data Library, July 1999
 A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome, stru
 A:Reference number: A75001
 A:Accession: E75148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-417 <KAM>
 A:Cross-references: GB:AJ248284; GB:AL096836; NID:q5457730; PIDN:CMD49268.1; PID:q545777
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB2116
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein Mj1154

Query Match 68.5%; Score 37; DB 2; Length 417;
 Best Local Similarity 72.7%; Pred. No. 28;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSOELAPIYEE 11
 |||||:
 Db 71 LSOELMAYEE 81

RESULT 9
 B59096
 Hypothetical protein PX01-42 - *Bacillus anthracis* virulence plasmid pX01
 C:Species: *Bacillus anthracis*
 C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
 C:Accession: B59096
 R:Okina, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler, J.; Bacteriol. 181, 6509-6515, 1999
 A:Title: Sequence and organization of pX01, the large *Bacillus anthracis* plasmid harbort
 A:Reference number: A59091; PMID:99445483; PMID:10515943
 A:Accession: B59096
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-452 <OKI>
 A:Cross-references: GB:AF065404; NID:q4894216; PIDN:AAD32346.1; PID:q4894258
 A:Experimental source: strain Sterne

A:Note: similar to hypothetical, glutamate rich protein; locus A45555; Plasmodium fa
 C:Genetics:
 A:Gene: PX01-42
 A:Genome: plasmid

Query Match 68.5%; Score 37; DB 2; Length 452;
 Best Local Similarity 63.6%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LSOELAPIYEE 11
 |||||:
 Db 219 LSOELAPIYEE 229

RESULT 10
 AB3130
 alpha-galactosidase [mel] (imported) - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 17-May-2002
 C:Accession: AB3130
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.H.; Chen, Y.; Wo
 etage, G.; Gilllet, W.; Grant, C.; Quentener, D.; Kutayavin, T.; Levy, R.; Li, M.; McC
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Petry, M.; Gordon-Ka
 sler, E.W.
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AB3130
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-457 <KUR>
 A:Cross-references: GB:AF008689; PIDN:AA145459.1; PID:q17743164; CSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: melA
 A:Map position: linear chromosome
 C:Superfamily: melIIBiose-specific alpha-galactosidase

Query Match 68.5%; Score 37; DB 2; Length 457;
 Best Local Similarity 63.6%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSOELAPIYEE 11
 |||||:
 Db 183 LAMHAPIYEE 193

RESULT 11
 D96157
 alpha-galactosidase [mel] (imported) - *Agrobacterium tumefaciens* (strain C58,
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002
 C:Accession: D96157
 R:Goodner, B.; Hinkle, G.; Gallun, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold
 A.; Liu, F.; Wolcott, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
 Science 294, 2323-2328, 2001
 A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*
 A:Reference number: A97359; PMID:11743194
 A:Accession: D96157
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-474 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK88782.1; PID:q5158531; CSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L436
 A:Map position: linear chromosome
 C:Superfamily: melIIBiose-specific alpha-galactosidase

Query Match 68.5%; Score 37; DB 2; Length 474;
 Best Local Similarity 63.6%; Pred. No. 33;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ISOEIAIPEE 11
 : : | | | | |
 Db 200 LAMHDIAPYEE 210

RESULT 12

JC7337
 thiazole biosynthetic enzyme - *Aspergillus oryzae*
 N:Alternate names: pyrithiamin resistance protein
 C:Species: *Aspergillus oryzae*
 C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 03-Aug-2001
 C:Accession: JC7337
 R:Kobayashi, T.; Yamashita, N.; Nishimura, A.
 Biosci. Biotechnol. Biochem. 64, 1416-1421, 2000
 A:Title: Pyrithiamin resistance gene (plrA) of *Aspergillus oryzae*: Cloning, characterization,
 A:Reference number: JC7337; MUID:20399355; PMID:10945258
 A:Accession: JC7337
 A:Molecule type: DNA
 A:Residues: 1-327 <KUB>
 A:Cross-references: GB:AF217503; PID:96694866; PIDN:AAV25444.1; PID:96694867
 A:Experiment: plrA source: strain wild type
 C:Genetics:
 A:Gene: plrA
 A:Introns: 113/1
 C:Superfamily: thiamin biosynthesis protein thl1

Query Match 66.7%; Score 36; DB 2; Length 327;
 Best Local Similarity 75.0%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 4 EIAIPEE 11
 : : | | | | |
 Db 138 ELGYPEE 145

RESULT 13

T34415
 hypothetical protein F07E5.9 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34415
 R:Du, Z.; Goela, D.
 Submitted to the EMBL Data Library, December 1996
 A:Description: The sequence of C. elegans cosmid F07E5.
 A:Reference number: 221520
 A:Accession: T34415
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-412 <DUZ>
 A:Cross-references: EMBL:U080837; PIDN:AA837908.1; GSPDH:GND0020; CUSP:F07E5.9
 A:Experimental source: strain Bristol N2; clone F07E5
 C:Genetics:
 A:Gene: CESP:F07E5.9
 A:Map position: 2
 A:Introns: 8/3; 34/3; 79/3; 125/1; 150/3; 179/2; 240/3; 343/3

Query Match 66.7%; Score 36; DB 2; Length 412;
 Best Local Similarity 63.6%; Pred. No. 44;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 1 ISOEIAIPEE 11
 : : | | | | |
 Db 7 ISPEIAVPEE 17

RESULT 14

A40597
 DNA-directed DNA polymerase (EC 2.7.7.7) I - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 17-Mar-2000
 C:Accession: A40597; E75363; S42112; S42235
 R:Gutman, P.D.; Fuchs, P.; Ouyang, L.; Minton, K.W.
 J. Bacteriol. 175, 3581-3590, 1993

A:Title: Identification, sequencing, and targeted mutagenesis of a DNA polymerase gene
 A:Reference number: A40597; MUID:93273728; PMID:8501062
 A:Accession: A40597

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-921 <GUT>
 A:Cross-references: EMBL:L14581; NID:929067; PIDN:AA36974.1; PID:929068
 A:Experimental source: strain R1
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.H.; Dodson, R.
 M.; Shen, M.; Vamshidevan, J.; Lam, P.; McDonald, L.; Ullrich, T.; Zaleski, C.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 266, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036996; PMID:10567266
 A:Accession: E75363
 A:Molecule type: DNA
 A:Residues: 1-MVPCGGGAGSCSIDFALCCLRGSGNYVSRITP, 1-12, 'P', 14-41, 'H', 43-72, 'F', 74-92
 A:Cross-references: GB:AF002012; GB:AF000513; NID:96459473; PIDN:AAV11264.1; PID:964
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRI707
 A:Map position: 1
 C:Superfamily: DNA-directed DNA polymerase I
 C:Keywords: nucleotidyltransferase

Query Match 66.7%; Score 36; DB 2; Length 921;
 Best Local Similarity 63.6%; Pred. No. 11e+02;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 1 ISOEIAIPEE 11
 : : | | | | |
 Db 768 LSNMGIPIYAR 778

RESULT 15

C86902
 hypothetical protein yxh [imported] - *Lactococcus lactis* subsp. *lactis* (strain 11.4)
 C:Species: *Lactococcus lactis* subsp. *lactis*
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: C86902
 R:Botolin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarne, K.; Weissenbach, J.; El
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*
 A:Reference number: A86625; MUID:2125186; PMID:11337471
 A:Accession: C86902
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-59 <SN>
 A:Cross-references: GB:AF005176; PID:912725285; PIDN:AAK06317.1; GSPDH:GND0146
 A:Experimental source: strain 11.403
 C:Genetics:
 A:Gene: yxh

Query Match 64.8%; Score 35; DB 2; Length 59;
 Best Local Similarity 77.8%; Pred. No. 7.7;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 ISOEIAIPEE 9
 : : | | | | |
 Db 10 LKOEIAIPEE 18

Search completed: January 15, 2003, 12:37:17
 Job time: 16.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 Seconds
(without alignments)
72.584 Million cell updates/sec

Title: US-09-823-649A-3
Sequence: 1 LSQELAIPIYEE 11

Scoring table: BIOSUM62
Gap: 10.0, Capext 0.5

Searched: 112892 segs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	832	1 DPOL_THERAQ	P18621 thermus agu
2	54	100.0	834	1 DPOL_THERCA	P80194 thermus cal
3	54	100.0	834	1 DPOL_THERTH	P52028 thermus the
4	44	81.5	831	1 DPOE_THERTH	P30313 thermus the
5	42	77.8	229	1 CG18_YEAST	P38794 saccharomyc
6	38	70.4	326	1 TH14_YEAST	P33318 saccharomyc
7	38	70.4	833	1 DPOL_THERPI	O52225 thermus fil
8	36	66.7	327	1 TH14_ASPOR	Q94429 aspergillus
9	36	66.7	956	1 VPOL_DEIRA	P52027 deinococcus
10	35	64.8	149	1 VEG_HPV31	P17386 human papil
11	35	64.8	1448	1 TRCS_ECOLI	P27190 escherichia
12	34	63.0	222	1 KCY_AQUAE	O67907 aquifex aeo
13	34	63.0	238	1 RM08_YEAST	P23353 saccharomyc
14	34	63.0	273	1 PAMB_PSEAE	O91363 pseudomonas
15	34	63.0	285	1 TPFB_YEAST	P08064 saccharomyc
16	34	63.0	342	1 GLK_ANASP	P58616 adiabena sp
17	34	63.0	382	1 NAWH_CLOPE	P10481 clostridium
18	34	63.0	390	1 MODA_SYNY3	O53368 synchocyst
19	34	63.0	415	1 CSD_HALNI	Q94441 halobacteri
20	34	63.0	499	1 C1W5_HUMAN	O95279 homo sapien
21	34	63.0	1136	1 POLG_OMV	P20234 o genome po
22	34	63.0	1324	1 POL2_GCMV	P13026 hungarian g
23	33	61.1	208	1 YL1J_ECOLI	P75805 escherichia
24	33	61.1	246	1 Y564_METTH	O26664 methanobact
25	33	61.1	274	1 Y110_HUMAN	Q15013 homo sapien
26	33	61.1	285	1 MPAF_CAEEL	Q09438 caenorhabdi
27	33	61.1	338	1 SP54_YEAST	P08937 saccharomyc
28	33	61.1	370	1 MYOM_APLCA	P15513 a mycomodul
29	33	61.1	413	1 YQ28_BACAN	Q94404 bacillus an
30	33	61.1	433	1 Y103_METUA	O57567 methanococc
31	33	61.1	455	1 SYS_PYROH	O58441 pyrococcus
32	33	61.1	487	1 SYS_CAEEL	Q18678 caenorhabdi
33	33	61.1	636	1 S6A7_HUMAN	Q99884 homo sapien

34	33	61.1	637	1 S6A7_RAT	P28573 rattus norv
35	32	59.3	140	1 VEG_HPV1A	P06929 human papil
36	32	59.3	227	1 ID11_HUMAN	Q13907 homo sapien
37	32	59.3	227	1 ID11_MESAU	O35586 mesocricetu
38	32	59.3	227	1 ID11_MOUSE	P58044 mus musculu
39	32	59.3	227	1 ID11_RAT	O35760 rattus norv
40	32	59.3	258	1 ARA5_ARATH	P28188 arabidopsis
41	32	59.3	266	1 TRN2_DARTST	P50163 datura sira
42	32	59.3	266	1 AKOH_HILPU	Q921X8 heliobacte
43	32	59.3	325	1 BMAP_BORGA	O31357 borrelia ga
44	32	59.3	340	1 RPOA_STRCO	P72404 streptomyce
45	32	59.3	340	1 RPOA_STRGT	Q944V6 streptomyce

ALIGNMENTS

AC	DPOL_THERAQ	STANDARD:	PRT:	832 AA.
AC	P19821:			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	DNA polymerase I, thermostable (EC 2.7.7.7) (Tag polymerase I).			
GN	POLA OR POLI.			
OS	Thermus aquaticus.			
OC	Bacteria; Thermus/deinococcus group; Deinococci; Thermales;			
OC	Thermaceae; Thermus.			
OX	NCBI_TaxID-271;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89197950; PubMed=2649500;			
RA	Lawyer F.C., Stoffel S., Sakai R.K., Myambo K., Drummond R.,			
RA	Celfand D.H.;			
RT	"Isolation, characterization, and expression in Escherichia coli of			
RT	the DNA polymerase gene from Thermus aquaticus.";			
RT	J. Biol. Chem. 264:6427-6437(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN-Y71;			
RX	MEDLINE=95204371; PubMed=7896728;			
RA	Ishino Y., Ueno T., Miyagi M., Demori T., Imamura M., Tsunawasa S.,			
RA	Kato I.;			
RT	"Overproduction of Thermus aquaticus DNA polymerase and its			
RT	structural analysis by ion-spray mass spectrometry.";			
RT	J. Biochem. 116:1019-1024(1994).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).			
RX	MEDLINE=95364959; PubMed=7637814;			
RA	Kim Y., Bom S.H., Wang J., Lee D.-S., Suh S.W., Steitz T.A.;			
RA	"Crystal structure of Thermus aquaticus DNA polymerase.";			
RA	Nature 376:612-616(1995).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 290-832.			
RX	MEDLINE=96016150; PubMed=7568114;			
RA	Korolev S., Nayal M., Barnes W.M., di Cera E., Waksman G.;			
RT	"Crystal structure of the large fragment of Thermus aquaticus DNA			
RT	polymerase I at 2.5-A resolution: structural basis for			
RT	thermostability.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 92:9264-9268(1995).			
RN	[5]			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE=96353982; PubMed=8717047;			
RA	Bom S.H., Wang J., Steitz T.A.;			
RT	"Structure of Tag polymerase with DNA at the polymerase active site.";			
RT	Nature 382:278-281(1996).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 295-832.			
RX	MEDLINE=99077817; PubMed=9857206;			
RA	Li Y., Korolev S., Waksman G.;			
RT	"Crystal structures of open and closed forms of binary and ternary			
RT	complexes of the large fragment of Thermus aquaticus DNA polymerase			

RT 1: structural basis for nucleotide incorporation.";
 RL EMBJ. J. 17:7514-7525(1998).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 290-832.
 RX MEDLINE=98266352; PubMed=9605316;
 RA Li Y., Kong Y., Korolev S., Maksman G.;
 RT "Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA
 polymerase I complexed with deoxyribonucleoside triphosphates.";
 RL Protein Sci. 7:1116-1123(1998).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=98445410; PubMed=9770525;
 RA Morali R., Sharkey D.J., Daise J.L., Murthy H.M.K.;
 RT "Crystal structure of Taq DNA polymerase in complex with an inhibitory
 Fab: the Fab is directed against an intermediate in the helix-coil
 dynamics of the enzyme.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12562-12567(1998).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.30 ANGSTROMS) OF 293-831.
 RX MEDLINE=99380545; PubMed=10449720;
 RA Li T., Mitaxov V., Maksman G.;
 RT "Structure-based design of Taq DNA polymerases with improved
 properties of dideoxynucleotide incorporation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9491-9496(1999).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + (DNA)(N).
 CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
 thermostability. Has a relatively high error rate because it lacks
 exonuclease proofreading functionality.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J04639; AAA27507.1; -;
 CC EMBL: D32013; BAA06775.1; -;
 CC PIR: A33530; A33530.
 DR PDB: 1TAQ; 07-DEC-96.
 DR PDB: 1TAU; 11-APR-97.
 DR PDB: 1KTQ; 08-NOV-96.
 DR PDB: 2KTQ; 22-DEC-99.
 DR PDB: 3KTQ; 22-DEC-99.
 DR PDB: 4KTQ; 22-DEC-99.
 DR PDB: 5KTQ; 30-SEP-98.
 DR PDB: 1PCK; 14-OCT-98.
 DR PDB: 1LXE; 12-MAR-97.
 DR PDB: 1QSS; 16-AUG-99.
 DR PDB: 1QSY; 16-AUG-99.
 DR PDB: 1QTM; 16-AUG-99.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR000513; Exo_N.1.
 DR InterPro: IPR003583; HHH.1.
 DR InterPro: IPR003584; HHH.2.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuc_N; 1.
 DR PRINTS: PR00868; DNAPOL1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00278; HhH1; 1.
 DR SMART: SM00279; HhH2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRfam: TIGR00593; pola; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 DR TRANSFERASE: DNA-directed DNA polymerase; DNA replication; DNA repair;
 KW DNA-binding; 3D-structure.
 FT DOMAIN 410 832 POLYMERASE (BY SIMILARITY).

FT CONFLICT 155 155 V -> A (IN REF. 2).
 SQ SEQUENCE 832 AA; 93909 MW; F1731055B5246F03 CRC64;
 Query Match 100.0%; Score 54; DB 1; Length 832;
 Best Local Similarity 100.0%; Fred. No. 0.011;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 LSQELAIPIYEE 11
 |||||||||
 Db 678 LSQELAIPIYEE 688
 RESURF 2
 ID DPOL_THECA STANDARD; PRT; 834 AA.
 AC P80194;
 DT 01-JUL-1993 (rel. 26, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE DNA polymerase I, thermostable (EC 2.7.7.7) (TAC polymerase 1).
 GN POLA.
 OS *Thermus caldophilus*.
 OC Bacteria; *Thermus*/Deinococcus group; Deinococci; Thermales;
 OC Thermaceae; *Thermus*.
 OX NCBI_TaxId=272;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GR24;
 RA Kwon S.-T., Kim J.S., Park J.H., Kim H., Lee D.-S.;
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-21.
 RC STRAIN=GR24;
 RX MEDLINE=93285135; PubMed=8508785;
 RA Park J.H., Kim J.S., Kwon S.-T., Lee D.-S.;
 RT "Purification and characterization of *Thermus caldophilus* GR24 DNA
 RT polymerase.";
 RL Eur. J. Biochem. 214:135-140(1993).
 CC -1- FUNCTION: THERMOSTABLE ENZYME THAT HAS 5'-TO-3' EXONUCLEASE
 CC ACTIVITY AND NO 3'-TO-5' EXONUCLEASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + (DNA)(N).
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U62584; AAB81398.1; -;
 CC PIR: S33287; S33287.
 DR HSP: P13821; 1TAQ.
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR001098; DNA_pol.
 DR InterPro: IPR002298; DNA_pol.
 DR InterPro: IPR000513; Exo_N.1.
 DR InterPro: IPR003583; HHH.1.
 DR InterPro: IPR003584; HHH.2.
 DR Pfam: PF00476; DNA_pol_A; 1.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuc_N; 1.
 DR PRINTS: PR00868; DNAPOL1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00278; HhH1; 1.
 DR SMART: SM00279; HhH2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRfam: TIGR00593; pola; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 DR TRANSFERASE: DNA-directed DNA polymerase; DNA replication; DNA repair;
 KW DNA-binding.

FT DOMAIN 412 834 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 834 AA; 93798 MW; A851F3C3076348E CRC64;

Query Match 100.0%; Score 54; DB 1; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LSQELAIPEE 11
Db 680 LSQELAIPEE 690

RESULT 3

ID DPROT_THERM STANDARD; PRT: 834 AA.
AC P502028;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (Tth polymerase I).
GN POLA.
OS Thermus thermophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=274;
RP SEQUENCE FROM N.A.
RC STRAIN-HB8 / ATCC 27634;
RA Asakura K., Komatsubara H., Soga S., Yomo T., Oka M., Eml S.,
RA Urabe I.;
RT "Cloning, nucleotide sequence, and expression in Escherichia coli of
RT DNA polymerase gene (polA) from Thermus thermophilus HB8.";
RL J. Ferment. Bioeng. 76:285-296(1993).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + (DNA)(N).
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL: D28878; BAA06033.1;
DR JISSP: P19821; 17A0.
DR InterPro: IPR002421; 5.3.exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N_1.
DR InterPro: IPR003583; HNH_1.
DR InterPro: IPR003584; HNH_2.
DR Pfam: PF00476; DNA_pol_A_1.
DR Pfam: PF01367; 5.3.exonuclease; 1.
DR Pfam: PF02739; 5.3.exonuc_N; 1.
DR PRINTS: PR00868; DNAPOL1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; HNH1; 1.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00482; POLAC; 1.
DR TIGRfams: TIGR00593; polA; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.
FT DOMAIN 412 834 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 834 AA; 94049 MW; 1A98145DC1A54A9 CRC64;

Query Match 100.0%; Score 54; DB 1; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LSQELAIPEE 11

Db 680 LSQELAIPEE 690

RESULT 4

ID DPROT_THERM STANDARD; PRT: 831 AA.
AC P30313;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (Tth polymerase I).
GN POLA OR POL.
OS Thermus thermophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=274;
RP SEQUENCE FROM N.A.
RC STRAIN-ACH H-1257;
RC MEDLINE-93087201; PubMed-1454544;
RA Akmetzjanov A.A., Vakhilov V.A.;
RT "Molecular cloning and nucleotide sequence of the DNA polymerase gene
RT from Thermus flavus.";
RL Nucleic Acids Res. 20:5839-5839(1992).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + (DNA)(N).
CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
CC thermostability.
CC -1- MISCELLANEOUS: Has an optimal temperature above 70 degrees
CC Celsius; can be used at temperatures up to 95 degrees Celsius.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL: X66105; CAA46900.1;
DR PIR: S24929; S24929.
DR PIR: S26675; S26675.
DR JISSP: P19821; 17A0.
DR InterPro: IPR002421; 5.3.exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N_1.
DR InterPro: IPR003583; HNH_1.
DR InterPro: IPR003584; HNH_2.
DR Pfam: PF00476; DNA_pol_A; 1.
DR Pfam: PF01367; 5.3.exonuclease; 1.
DR Pfam: PF02739; 5.3.exonuc_N; 1.
DR PRINTS: PR00868; DNAPOL1.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; HNH1; 1.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00482; POLAC; 1.
DR TIGRfams: TIGR00593; polA; 1.
DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.
FT DOMAIN 409 831 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 831 AA; 93783 MW; 96F93CEFA3CA536D CRC64;

Query Match 81.5%; Score 44; DB 1; Length 831;
Best Local Similarity 81.8%; Pred. No. 0.95;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 LSQELAIPEE 11

Db 677 LSGELSLPYEE 687

RESULT 5

CG18 YEAST

STANDARD;

PRT: 229 AA.

ID CG18 YEAST

AC P38794;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE G1/S-specific cyclin PCL5.

GN PCL5 OR YHR071W

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RX MEDLINE=94376003; PubMed=8091229;

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,

RA Du Z., Favello A., Fulton L., Gattling S., Geisel C., Kirsten J.,

RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,

RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,

RA Nhan M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,

RA Vignati D., Wilcox L., Woldman P., Waterston R., Wilson R.,

RA Vaudin M.;

RT Complete nucleotide sequence of Saccharomyces cerevisiae chromosome

VIII.

RT Science 265:2077-2082(1994).

CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. MOST SIMILAR TO G1/S-

SPECIFIC CYCLINS PCLT AND PCL2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>CC or send an email to license@isb-sib.ch).

CC

DR EMBL: U00061; AAB68375.1.

DR PIR: S46696; S46696.

DR SGD: S000113; PCL5.

DR InterPro: IPR004366; Cyclin.

DR Pfam: PF00134; cyclin; 1.

DR SMART: SM00385; CYCLIN; 1.

DR PROSITE: PS00292; CYCLIN; FALSE_NEG.

DR Cyclin; cell cycle; cell division.

KW SHODNCE 229 AA; 26467 MW; 86839ABBD5DD49 CRC64;

SQ

Query Match

Best Local Similarity 77.8%; Score 42; DB 1; Length 229;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY

1 LSGELSLPYEE 11

Db

174 LNEELALPYDE 184

RESULT 6

TH14 YEAST

ID TH14 YEAST

AC P32318;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Thiazole biosynthetic enzyme, mitochondrial precursor.

GN TH14 OR M011 OR ESP35 OR YGR144W OR G6620.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RX MEDLINE=96158062; PubMed=8585325;

RA Skala J., Nawrocki A., Goffeau A.;

RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Valla E.,

RA Nawrocki A., del Bino S., Goffeau A.;

RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 205-326 FROM N.A.

RC STRAIN=S288C;

RX MEDLINE=96158062; PubMed=8585325;

RA Skala J., Nawrocki A., Goffeau A.;

RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Valla E.,

RA Nawrocki A., del Bino S., Goffeau A.;

RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [14]

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93070608; PubMed=1441749;

RA Praekelt U.M., Meacock P.A.;

RT "MOL1, a Saccharomyces cerevisiae gene that is highly expressed in

early stationary phase during growth on molasses.";

RT Yeast 8:699-710(1992).

RN [2]

RP SEQUENCE FROM N.A.

RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Valla E.,

RA Nawrocki A., del Bino S., Goffeau A.;

RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 205-326 FROM N.A.

RC STRAIN=S288C;

RX MEDLINE=96158062; PubMed=8585325;

RA Skala J., Nawrocki A., Goffeau A.;

RT "The sequence of a 27 kb segment on the right arm of chromosome VII

from Saccharomyces cerevisiae reveals MOL1, NAT2, RPL30B, RSK1, CTS4,

PW1/CHO2, NSR1 genes and ten new open reading frames.";

RT Yeast 11:1421-1427(1995).

RN [4]

RP THIAMINE REGULATION.

RX MEDLINE=95028146; PubMed=7941734;

RA Praekelt U.M., Byrne K.L., Meacock P.A.;

RT "Regulation of TH14 (MOL1), a thiamine-biosynthetic gene of

Saccharomyces cerevisiae.";

RT Yeast 10:481-490(1994).

RN [5]

RP FUNCTION IN DNA DAMAGE TOLERANCE.

RX MEDLINE=98035046; PubMed=9367751;

RA Machado C.R., Praekelt U.M., de Oliveira R.C., Barbosa A.C.,

RA Byrne K.L., Meacock P.A., Menck C.F.;

RT "Dual role for the yeast TH14 gene in thiamine biosynthesis and DNA

damage tolerance.";

RT J. Mol. Biol. 273:114-121(1997).

CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF THE THIAMINE PRECURSOR

THIAOLE. ALSO SEEMS TO HAVE A ROLE IN MITOCHONDRIAL DNA DAMAGE

TOLERANCE.

CC -1- SUBCELLULAR LOCATION: Mitochondrial (potential).

CC -1- INDUCTION: RERESSED BY THIAMINE.

CC -1- MISCELLANEOUS: EXPRESSED AT HIGH LEVELS IN THE EARLY STATIONARY

PHASE OF BATCH CULTURES GROWING ON MOLASSES, AN INDUSTRIAL MEDIUM.

CC -1- SIMILARITY: BELONGS TO THE TH14 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL: X61669; CAA3343.1;

DR EMBL: X65807; CAA59802.1;

DR EMBL: E72929; CAA97157.1;

DR EMBL: E72930; CAA97159.1;

DR PIR: S17019; S17019.

DR PIR: S25321; S25321.

DR SGD: S0003376; TH14.

DR InterPro: IPR002922; TH14.

DR Pfam: PF01946; TH14; 1.

DR TIGRfam: TIGR00292; TH14; 1.

KT Thiamine biosynthesis; Mitochondrion; Transit peptide; PAD; NAD.

FT TRANSIT 1

FT CHAIN ? 326

FT NP BIND ? 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

FT NP BIND 67 97

Query Match

Best Local Similarity 70.4%; Score 38; DB 1; Length 326;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;


```

RT Deinococcus radiodurans."
RL J. Bacteriol. 175:3581-3590(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofatt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
CC 5' MILARITY').
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + (DNA)(N).
CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L14581; AAC36974.1;
DR EMBL: AE002012; AAF1264.1;
DR HSSP: P19821; ZKTQ.
DR TIGR: D81707;
DR InterPro: IPR002421; S_3_exonuclease.
DR InterPro: IPR001098; DNA_Pol.
DR InterPro: IPR002298; DNA_Pol.
DR InterPro: IPR00513; EXO_N.1.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR003584; HHH_2.
DR Pfam: PR00476; DNA_Pol_A.1.
DR Pfam: PR01367; S_3_exonuclease.1.
DR Pfam: PR02739; S_3_exonuclease.1.
DR PRINTS: PR00868; DNAPOLI.
DR SMART: SM00475; 53EXOC.1.
DR SMART: SM00278; Hhh1.1.
DR SMART: SM00279; Hhh2.1.
DR SMART: SM00482; POLAC.1.
DR TIGRfam: TIGR00593; pola.1.
DR PROSITE: PS00447; DNA_POLYMERASE_A.1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW Hydrolyase; Exonuclease; DNA-binding; Complete proteome.
FT CONFLICT 48 48 T -> A (IN REF. 1).
FT CONFLICT 77 77 H -> D (IN REF. 1).
FT CONFLICT 108 108 F -> L (IN REF. 1).
FT CONFLICT 128 128 P -> R (IN REF. 1).
FT CONFLICT 406 406 P -> R (IN REF. 1).
FT CONFLICT 540 540 L -> V (IN REF. 1).
SQ SEQUENCE 956 AA; 105659 MW; 6ABBP117D5AB84A CRC64;

Query Match 66.7%; Score 36; DB 1; Length 956;
Best Local Similarity 63.6%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

ID V06_HPV31 STANDARD; PRT; 149 AA.
AC P17386;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E6 protein.
GN E6.
OS Human papillomavirus type 31.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6929478; PubMed=2545036;
RA Goldsborough M.D., Disilverie D., Temple G.F., Lorincz A.T.;
RT "Nucleotide sequence of human papillomavirus type 31: a cervical
RT neoplasia-associated virus."
RL Virology 171:306-311(1989).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL
CC OF THIS VIRUS (CERVICAL NEOPLASIA-ASSOCIATED VIRUS).
CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
CC STRANDED DNA (IN VITRO).
CC -1- SUBCELLULAR LOCATION: Nuclear matrix-associated.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J04353; AAA46950.1;
DR PIR: A32444; M6M131.
DR InterPro: IPR001334; E6.
DR Pfam: PF00518; E6.1.
KW Early protein; DNA-binding; Nuclear protein; zinc-finger;
KW Oncogene.
FT ZN_FING 30 66 POTENTIAL.
FT ZN_FING 103 139 POTENTIAL.
SQ SEQUENCE 149 AA; 17713 MW; 61D2A86C362767D9 CRC64;

Query Match 64.8%; Score 35; DB 1; Length 149;
Best Local Similarity 63.6%; Pred. No. 9.1;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

RESULT 10
V06_HPV31

RESULT 11
TRCS_ECOLI STANDARD; PRT; 1448 AA.
ID TRCS_ECOLI STANDARD; PRT; 1448 AA.
AC P27190; P27191; P27185.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA primase (rac (EC 2.7.7.-) (Replication primase).
GN TRAC.
OS Escherichia coli.
OC Plasmid Incp-beta R751.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 2-11; 219-234 AND 702-714.
RC STRAIN-H101;
RX MEDLINE=92297959; PubMed=1818755;
RA Miele L., Strack B., Kruff V., Janka E.;
RT "Gene organization and nucleotide sequence of the primase region of
RT Incp plasmids RP4 and R751.";
RL DNA Seq. 2:145-162(1991).

```

RN [2]
RP SEQUENCE FROM N.A.
RA Thomas C.M.;
RL Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR AUTONOMOUS REPLICATION IN ESCHERICHIA COLI.
CC TRANSFERRED INTO THE RECIPIENT CELL DURING BACTERIAL CONJUGATION.
CC CATALYZES THE SYNTHESIS OF SHORT OLIGORIBONUCLEOTIDE PRIMERS WITH
CC CPA OR PCPA AT THEIR 5'-TERMINI ON A SINGLE STRANDED TEMPLATE DNA.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; TRAC-2 (SHOWN HERE), TRAC-3 AND
CC TRAC-4; ARE PRODUCED BY ALTERNATIVE INITIATION.
CC -1- SIMILARITY: TO PLASMID INCP-ALPHA RP4 TRAC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X59794; CAA42460.1; -
DR EMBL: U67184; AAC64468.1; -
DR PIR: S37669; S37669.
DR InterPro: IPR002936; DNAPrim_toprim.
DR Pfam: PF01751; Toprim; 1.
DR SMART: SM00493; TOPRIM; 1.
DR DNA replication: Transferase; DNA-directed RNA polymerase; Plasmid;
KM Alternative Initiation.
RN CHAIN 2 1448 DNA PRIMASE, ISOFORM TRAC-2.
FT CHAIN 219 1448 DNA PRIMASE, ISOFORM TRAC-3.
FT CHAIN 702 1448 DNA PRIMASE, ISOFORM TRAC-4.
FT INT_MET 1 1 FOR ISOFORM TRAC-2.
FT INT_MET 219 219 FOR ISOFORM TRAC-3.
FT INT_MET 702 702 FOR ISOFORM TRAC-4.
FT INT_MET 1448 AA; 158950 MW; 1C0BAE67C00D4E04 CRC64;
SQ SEQUENCE

Query Match 64.8%; Score 35; DB 1; Length 1448;
Best Local Similarity 66.7%; Pred. NO. 97;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 DELTAIPYEE 11
DB 986 QYLAIPYEQ 994

RESULT 12
KEY AOUAE STANDARD; PRI: 222 AA.
ID KEY AOUAE STANDARD; PRI: 222 AA.
AC 067907;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytidylate kinase (EC 2.7.4.14) (CK) (Cytidine monophosphate kinase)
DE (CMP kinase).
DE CMK OR AO 2153.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE-98196666; PubMed-9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Jenox A.L.,
RA Feldman D.E., Overbeek R., Shead M.A., Keller M., Aubay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -1- CATALYTIC ACTIVITY: AMP + (d)CMP -> ADP + (d)CDP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE CYTIDYLATE KINASE FAMILY. SUBFAMILY 1.
CC

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE000775; AAC07867.1; -
DR HSSP: P23863; 1CKE.
DR InterPro: IPR003136; Cytidylate_kin.
DR Pfam: PF02224; Cytidylate_kin; 1.
DR TIGRFAMs: TIGR00017; cmk; 1.
KM Transferase; Kinase; ATP-binding; Complete proteome.
FT NP_BIND 7 15 AMP (BY SIMILARITY).
SQ SEQUENCE 222 AA; 25400 MW; 523BDCDFB48128C8 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 222;
Best Local Similarity 63.6%; Pred. NO. 22;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ISQRIALPYEE 11
DB 20 ISQKINIPYLE 30

RESULT 13
KEY YEAST STANDARD; PRI: 228 AA.
ID KEY YEAST STANDARD; PRI: 228 AA.
AC P22353;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Mitochondrial 60S ribosomal protein L8 (Yml8).
DE MRPL8 OR YJL063C OR J1125 OR HRD238.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-07173;
RX MEDLINE-90221879; PubMed-2183197;
RA Kitakawa M., Grohmann L., Grack H.-R., Isono K.;
RT "Cloning and characterization of nuclear genes for two mitochondrial
RT ribosomal proteins in Saccharomyces cerevisiae."
RL Nucleic Acids Res. 18:1521-1529(1990).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-S288c;
RX MEDLINE-95282514; PubMed-7762402;
RA Vandenbol M., Durand P., Dion C., Portetelle D., Huger F.;
RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
RT Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
RT L8."
RL Yeast 11:57-60(1995).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE L17P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X53841; CAA37834.1; -
DR EMBL: Z34286; CAA84060.1; -
DR EMBL: Z49338; CAA89354.1; -
DR PIR: S14890; S14890.
DR PIR: S47128; S47128.
DR SGD: S0003599; MRPL8.

```


Initiation factor; Protein biosynthesis; Zinc-finger.
KW DOMAIN 16 23 LYS-RICH (BASIC).
FT DOMAIN 49 56 LYS-RICH (BASIC).
FT DOMAIN 82 89 LYS-RICH (BASIC).
FT ZN-FING 236 262 CA-TYPE (POTENTIAL).
SQ SEQUENCE 285 AA; 31574 MW; 34DE9323876CEB95 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 285;
Best Local Similarity 45.5%; Pred. No. 28;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSOELAIPEE 11
: 11: 11 1
DB 123 IQQVGLPYSE 133

Search completed: January 15, 2003, 11:21:21
Job time : 8.28572 secs

Best Local Similarity	77.8%	Pred. No. 27;
Matches 7; Conservative	2;	Mismatches 0; Indels 0; Gaps 0

QY 1 LSGELAIPYEE 9
 DB 346 LSGELAIPYEE 354

RESULT 2

ID 066691 PRELIMINARY; PRT: 850 AA.
 AC 066691;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE Hypothetical protein AQ_367.
 GN AQ_367.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBL_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE:98196666; PubMed-9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,
 Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus".
 RT Nature 392:353-358(1998).
 RL EMBL: AE000686; AAC06556.1; -
 DR InterPro: IPR001633; EAL.
 DR InterPro: IPR00160; GGDEF.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00563; EAL; 1.
 DR Pfam: PF00990; GGDEF; 1.
 DR SMART: SM00267; DUF1; 1.
 DR SMART: SM00352; DUF2; 1.
 DR TIGRPFAM: TIGR00254; GGDEF; 1.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 850 AA; 97669 MW; C02DCEP91DNCDB0 CRC64;

Query Match 72.2%; Score 39; DH 16; Length 850;
 Best Local Similarity 54.5%; Pred. NO. 45;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSGELAIPYEE 11
 DB 227 SRMDVPEE 237

RESULT 3

ID 09X148 PRELIMINARY; PRT: 1451 AA.
 AC 09X148;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE P9L1.15 protein.
 GN P9L1.15.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBL_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Vrotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
 Lee J., Liu A., Li J., Kremenetskaia I., Luoro J., Gonzalez A.,
 Altai H., Araujo R., Brooks S., Buehler E., Chao O., Conn L.,
 Conway A.B., Dunn P., Hansen N., Huizar L., Khan S., Kim C., Palm C.,
 Rowley D., Shin P., Walker M., Davis R.W., Ecker J.R.,
 Federspiel N.A., Theologis A.;

RT "Arabidopsis thaliana chromosome 1 BAC P9L1 sequence."
 DR Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AC007591; RAD34650.1; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PP000006; ABC_transport; 1.
 DR SMART: SM00382; AAA; 1.
 KW ATP-binding.
 SQ SEQUENCE 1451 AA; 163602 MW; 0D35414D9C370A85 CRC64;

Query Match 70.4%; Score 38; DH 10; Length 1451;
 Best Local Similarity 54.5%; Pred. NO. 12e+02;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

RESULT 4

ID 09Y171 PRELIMINARY; PRT: 417 AA.
 AC 09Y171;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE Hypothetical protein PAB2116.
 GN PAB2116.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBL_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RA Helling R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 structure and evolution."
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248284; CAB49268.1; -
 DR InterPro: IPR002819; HD; 1.
 DR InterPro: IPR003607; MB_Ppase_HDC.
 DR Pfam: PF01966; HD; 1.
 DR SMART: SM00471; HDC; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 417 AA; 47906 MW; 59C62CB32C441833 CRC64;

Query Match 68.5%; Score 37; DH 17; Length 417;
 Best Local Similarity 72.7%; Pred. NO. 54;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSGELAIPYEE 11
 DB 71 LSGELAIPYEE 81

RESULT 5

ID 09X312 PRELIMINARY; PRT: 452 AA.
 AC 09X312;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE PX01-42.
 OS Bacillus anthracis.
 OC Plasmid virulence plasmid PX01.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBL_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-STERNK;


```

RT Investigating biology. The C. elegans Sequencing Consortium.
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Du Z., Goela D.;
RT "The sequence of C. elegans cosmid F0755 ";
RU Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RU Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80837; AAB37908.1;
DR InterPro; IPR002086; Aldohyde_dehydr.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 412 AA; 46671 MW; 8C8ACE5E416A294 CRC64;

Query Match
Best Local Similarity 66.7%; Score 36; DB 5; Length 412;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
Db 7 LSQELAIPEE 17

RESULT 9
O8SVP8 PRELIMINARY; PRT; 627 AA.
AC O8SVP8;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein EC004_1610.
GN EC004_1610.
OS Eucephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryontidae; Eucephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-GB-M1;
RX MEDLINE-21576510; PubMed-11719806;
RA Kellina M.D., Duprat S., Cornillot E., Metenier G., Thomaret P.,
RA Prenier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RA Delbec F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissbach J., Vivas C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Eucephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590444; CAD5350.1;
KW Initiation factor; Hypothetical protein.
SQ SEQUENCE 627 AA; 72118 MW; B8C3786244B820D CRC64;

Query Match
Best Local Similarity 66.7%; Score 36; DB 5; Length 627;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPEE 11
Db 472 LSQELAIPEE 482

RESULT 10
O9CDK0 PRELIMINARY; PRT; 59 AA.
ID O9CDK0

```

```

AC O9CDK0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Unknown protein.
GN YXAB OR LL2219.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IL1403;
RX MEDLINE-21235186; PubMed-11337471;
RA Bolotin A., Wincker P., Mager S., Jallion O., Matarne K.,
RA Weissbach J., Wirthlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006450; AKK06317.1;
KW Complete proteome.
SQ SEQUENCE 59 AA; 7043 MW; 80903DC52808162A CRC64;

Query Match
Best Local Similarity 64.8%; Score 35; DB 16; Length 59;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPE 9
Db 10 LSQELAIPE 18

RESULT 11
O9JEU9 PRELIMINARY; PRT; 95 AA.
ID O9JEU9
AC O9JEU9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Integrase (Fragment).
GN POL.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CMG30;
RX MEDLINE-20129484; PubMed-10664385;
RA Moudjeka I., Bikonon B., Zekenn L., Takehisa J., Harada Y.,
RA Yamaguchi-Kanda Y., Taniuchi Y., Ido E., Kaplue L., Muelle P.,
RA Parra H.O., Ikeda M., Hayami M., Miura T.;
RT "Genetic diversity of HIV-1 group M from Cameroon and Republic of
RT Congo.";
RL Arch. Virol. 144:2291-2311(1999).
DR EMBL; AF120196; AAF68138.1;
DR InterPro; IPR001584; Rve.
DR pfam; PF00665; rve; 1.
FT NON-TER 1
FT NON-TER 95
SQ SEQUENCE 95 AA; 10490 MW; AFB8FD5EC6765D32 CRC64;

Query Match
Best Local Similarity 64.8%; Score 35; DB 15; Length 95;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPE 9
Db 47 IAQELGIPY 55

RESULT 12
O9CKO5 PRELIMINARY; PRT; 202 AA.
ID O9CKO5

```

AC 09C05;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein PM1551.
 GN PM1551.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.-J., Zhang Q., Li L.-L., Paustian M.L., Whitlam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL; AF006192; AAK03635.1; -;
 DR InterPro: IPR004046; GST_Cterm.
 DR InterPro: IPR004045; GST_Nterm.
 DR Pfam; PF02798; GST_N.1.
 KW Hypothetical protein, complete proteome.
 SQ SEQUENCE 202 AA; 23180 MW; 5E6D95AE3E02E73 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 202;
 Best Local Similarity 70.0%; Pred. No. 63;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPE 10
 I :|||||:I
 Db 20 LKELAIPE 29

RESULT 13
 ID 09A3W3 PRELIMINARY; PRT; 208 AA.
 AC 09A3W3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Glutathione S-transferase family protein.
 GN CC3088.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohia N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Debey R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty K.,
 RA Utterberg S., Tran K., Wolf A., Yanathien J., Emoliava M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005972; AAK25050.1; -;
 DR TIGR; CC3088; -;
 DR InterPro: IPR004046; GST_Cterm.
 DR InterPro: IPR004045; GST_Nterm.
 DR Pfam; PF00043; GST_C.1.
 DR Pfam; PF02798; GST_N.1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 208 AA; 22885 MW; A48AC2988C284DD1 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 208;
 Best Local Similarity 60.0%; Pred. No. 65;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSQELAIPE 10

Db 18 LKELSLPE 27
 I :||:||||
 RESULT 14
 ID 099X18 PRELIMINARY; PRT; 240 AA.
 AC 099X18;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein SAV0203.
 GN SAV0203 OR SA0197.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=158878; 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50) and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohia T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsunari H., Maruyama A., Murakami H., Hosooyama A.,
 RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hiraoka H., Kubara S., Goto S., Yabuzaki J.,
 RA Kenehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003358; BAB56365.1; -;
 DR EMBL; AP003128; BAB1418.1; -;
 DR InterPro: IPR003593; AAA_Atpase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 KW Hypothetical protein, complete proteome.
 SQ SEQUENCE 240 AA; 27319 MW; 29C122944033DFAE CRC64;

Query Match 64.8%; Score 35; DB 16; Length 240;
 Best Local Similarity 70.0%; Pred. No. 75;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 SQELAIPE 11
 I :||:||||
 Db 230 SQKVIPE 239

RESULT 15
 ID 08RB89 PRELIMINARY; PRT; 291 AA.
 AC 08RB89;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate
 lyase.
 GN LYE OR TFE0933.
 OS Thermobacterium tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermobacteriales; Thermobacteriaceae; Thermobacter.
 OX NCBI_TaxID=19072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB47 / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 Tan H., Chen R., Wang J., Yu J., Yang H.;

RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013059; AAM24189.1; -
KW Transferase; Aminotransferase; Complete proteome.
SQ SEQUENCE 291 AA; 32346 MW; 105072D27AB06FEE CRC64;

Query Match 64.8%; Score 35; DB 16; Length 291;
Best Local Similarity 54.5%; Pred. No. 92;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSOELAIPEE 11
::|||::|||
DB 218 IARELNIPPEE 228

Search completed: January 15, 2003, 12:35:45
Job time : 32.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 11:21:44 ; Search time 130.143 Seconds
(without alignments)
190.345 Million cell updates/sec

Title: US-09-823-649a-3
Perfect score: 54
Sequence: 1 ISODELALPYEE 11

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODEL=framed_p2n.model -DEV=xlh
-O=us09823649.rng -FUNAT_14012003_151001_29119/app.query.fasta.1.1393
-DB=N.Geneseq_101002 -OFMT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -STARF=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.qcgn.1.1.0 @runat_14012003_151001_29119 -NCPU=6 -TCPU=3
-NO_XLPYX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMBOUT-120
-NARN.TIMBOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N.Geneseq_101002: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	54	100.0	1600	16	AA080745	5' Nuclease from T
2	54	100.0	1600	17	AA127686	Mutant Thermus aqu
3	54	100.0	1600	18	AA170347	Synthetis deficient
4	54	100.0	1600	18	AA176647	Tag gene 5' nuclea
5	54	100.0	1600	19	AA165786	Thermus aquaticus
6	54	100.0	1600	20	AA153855	Nucleotide sequenc
7	54	100.0	1600	19	AA163407	DNA sequence of a
8	54	100.0	1635	13	AA163407	Mutant thermostabl
9	54	100.0	1635	13	AA163407	Mutant thermostabl
10	54	100.0	1635	13	AA163407	Mutant thermostabl
11	54	100.0	1682	21	AA179559	Thermotable DNA d
12	54	100.0	1682	18	AA179559	Mutant Tag polymer
13	54	100.0	1682	18	AA179559	Mutant Tag polymer
14	54	100.0	1682	20	AA179559	Mutant Tag polymer
15	54	100.0	1682	20	AA179559	Mutant Tag polymer
16	54	100.0	1682	20	AA179559	Mutant Tag polymer
17	54	100.0	1682	20	AA179559	Mutant Tag polymer
18	54	100.0	1682	21	AA179559	Mutant Tag polymer
19	54	100.0	1682	21	AA179559	Mutant Tag polymer
20	54	100.0	1682	21	AA179559	Mutant Tag polymer
21	54	100.0	1682	21	AA179559	Mutant Tag polymer
22	54	100.0	1682	21	AA179559	Mutant Tag polymer
23	54	100.0	1682	21	AA179559	Mutant Tag polymer
24	54	100.0	1682	21	AA179559	Mutant Tag polymer
25	54	100.0	1682	21	AA179559	Mutant Tag polymer
26	54	100.0	1682	21	AA179559	Mutant Tag polymer
27	54	100.0	1682	21	AA179559	Mutant Tag polymer
28	54	100.0	1682	21	AA179559	Mutant Tag polymer
29	54	100.0	1682	21	AA179559	Mutant Tag polymer
30	54	100.0	1682	21	AA179559	Mutant Tag polymer
31	54	100.0	1682	21	AA179559	Mutant Tag polymer
32	54	100.0	1682	21	AA179559	Mutant Tag polymer
33	54	100.0	1682	21	AA179559	Mutant Tag polymer
34	54	100.0	1682	21	AA179559	Mutant Tag polymer
35	54	100.0	1682	21	AA179559	Mutant Tag polymer
36	54	100.0	1682	21	AA179559	Mutant Tag polymer
37	54	100.0	1682	21	AA179559	Mutant Tag polymer
38	54	100.0	1682	21	AA179559	Mutant Tag polymer
39	54	100.0	1682	21	AA179559	Mutant Tag polymer
40	54	100.0	1682	21	AA179559	Mutant Tag polymer
41	54	100.0	1682	21	AA179559	Mutant Tag polymer
42	54	100.0	1682	21	AA179559	Mutant Tag polymer
43	54	100.0	1682	21	AA179559	Mutant Tag polymer
44	54	100.0	1682	21	AA179559	Mutant Tag polymer
45	54	100.0	1682	21	AA179559	Mutant Tag polymer

ALIGNMENTS

RESULT 1
ID AA080745 standard; DNA: 1600 bp.

AC AA080745;
DT 19-JUL-1995 (first entry)
DE 5' Nuclease from Tag DNAP.
KW DNA-pol;metrass; DNAP; Tag; DNA cleavage; RNA cleavage; 5' nuclease; ds.
XX Thermus aquaticus YT-1.
OS Thermus aquaticus YT-1.
PN WO9429482-A.
XX
PD 22-DEC-1994.
XX

PF 06-JUN-1994; 94MO-US06253.
 XX
 PR 04-JUN-1993; 93US-0073384.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Brow MAD, Dahlberg JE, Lyanichev VI;
 XX
 DR WPI; 1995-036504/05.
 XX
 PT 5' Nuclease(s) derived from thermostable DNA polymerase(s) - have
 PT cleavage activity with reduced synthetic ability, used for
 PT detection of specific target sequences.
 XX
 PS Claim 5; Page 95-96; 159pp; English.
 XX
 CC The Tag DNA-polymerase (DNAP) gene was amplified by PCR. Amplified
 CC fragments were ligated into pMTQ18 vector, which contains the
 CC hybrid trp-lac (tac) promoter. An amplification/selection
 CC protocol was used to isolate clone 4B containing a mutated Tag DNAP
 CC gene (muttag) (sequence given in AA080745) having normal 5' nuclease
 CC activity but less than 1% of the wt tag DNAP activity. muttag was
 CC cut from pMTQ18 by EcoRI-SalI digestion and cloned into pET-3c.
 CC Thus clone was digested with BstXI and BamHI. The DNA was treated
 CC with DNAPcI Klenow fragment and dNTPs, blunt-ended and religated,
 CC resulting in an in-frame deletion of 903 nucleotides. The DNA
 CC sequence of the resulting 5' nuclease is given in AA080745.
 XX
 SQ Sequence 1600 BP; 287 A; 518 C; 553 G; 242 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0649 Length: 1600
 Score: 54.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 16 Gaps: 0
 US-09-823-649A-3 (1-11) x AA080745 (1-1600)
 QY 1 LeuSerGlnGluAlaAlaIleProTyrGluGlu 11
 |||||
 DB 1133 CTTCTCCAGCAGCTACGCTACCTCCCTACGAGAG 1165
 RESULT 2
 ID AAT27686 standard; DNA; 1600 BP.
 XX
 AC AAT27686;
 XX
 DT 18-NOV-1996 (first entry)
 XX
 DE Mutant Thermus aquaticus DNA polymerase coding sequence (Clone 4F).
 XX
 KW p53; mutant; mutation; cleavage; nuclease; cleavage; Thermus;
 KW Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
 KW Staphylococcus; Identification; detection; ds.
 XX
 OS Thermus aquaticus.
 XX
 FT Key Location/Qualifiers
 FT CDS 14..1600
 FT /tag= a
 FT /product= Mutant DNA polymerase.
 XX
 PN MO9615267-A1.
 XX
 PD 23-MAY-1996.
 XX
 PF 09-NOV-1995; 95MO-US14673.
 XX
 PF 30-AUG-1995; 95US-0520946.
 XX
 PR 09-NOV-1994; 94US-0337164.
 XX

PR 09-MAR-1995; 95US-0402601.
 XX
 PR 07-JUN-1995; 95US-0484956.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Brow MAD, Dahlberg JE, Fors L, Heisler LM, Lyanichev VI;
 XX
 DR Oldenburg MC, Olive DM;
 XX
 DR WPI; 1996-259862/26.
 XX
 PT P-PSDB: AAR96267.
 XX
 PT Cleavage of nucleic acids to detect mutation(s) - allows detection
 PT esp. in human p53 gene, to identify strains of microorganisms and
 PT viruses
 XX
 PS Example 2A; Page 257-258; 433pp; English.
 XX
 CC Cleavage of nucleic acids using an enzyme, especially a nuclease
 CC selected from the group consisting of: cleavage (RTM) RN enzyme,
 CC Thermus aquaticus DNA polymerase, "Thermus thermophilus DNA
 CC polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
 CC Rad1/Rad10 complex. The nucleic acid substrate is preferably an
 CC oligonucleotide containing a human p53 gene sequence or
 CC alternatively, microbial gene sequences. Cleavage products are
 CC compared to the cleavage products of reference gene sequences. The
 CC method is used for detecting mutation in the human p53 gene; for
 CC identifying strains of microorganisms, especially bacteria selected
 CC from the group of members of the genera Campylobacter,
 CC Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
 CC The method may also be used for the identification of viruses,
 CC especially hepatitis C virus and simian immunodeficiency virus.
 CC Thermus aquaticus (Tag) DNA polymerase was amplified using two
 CC primers (AAT27679, AAT27680). The Tag polymerase DNA was inserted into
 CC the BamHI restriction site of the expression vector pET-3c and mutant
 CC genes were created from that construct. This mutant was created
 CC after the vector was digested with BstXI and BamHI. The DNA was
 CC then treated with the Klenow fragment of DNAPcI to trim both 3'
 CC overhangs to blunt ends which were then ligated together, resulting
 CC in an in frame deletion of 903 nucleotides. The resulting mutant
 CC Tag polymerase is also referred to as the cleavage BB enzyme.
 XX
 SQ Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0649 Length: 1600
 Score: 54.00 Matches: 11
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 17 Gaps: 0
 US-09-823-649A-3 (1-11) x AAT27686 (1-1600)
 QY 1 LeuSerGlnGluAlaAlaIleProTyrGluGlu 11
 |||||
 DB 1133 CTTCTCCAGCAGCTACGCTACCTCCCTACGAGAG 1165
 RESULT 3
 ID AAT70347 standard; DNA; 1600 BP.
 XX
 AC AAT70347;
 XX
 DT 03-APR-1998 (first entry)
 XX
 DE Synthesis deficient Tag DNA polymerase gene clone 4F.
 XX
 KW DNA polymerase; tag polymerase gene; DNAP; 5' nuclease activity;
 KW cleavage BB; DNA cleavage; reduced synthetic activity; ss.
 XX
 OS Synthetic.
 XX
 OS Thermus aquaticus.
 XX

```

OS Thermus species sps17.
XX Key Location/Qualifiers
XX old_sequence 1..2
FT /tag= a
FT /note= "nucleotides 4-453 deleted from the native
FT sequence."
XX
XX M09206200-A.
XX
XX 16-APR-1992.
XX
XX 30-SEP-1991; 91WO-US07035.
XX
XX 28-SEP-1991; 90US-0590213.
XX 28-SEP-1991; 90US-0590466.
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU ) CETUS CORP.
XX
XX Abramson RD, Gelfand DH;
XX
XX WPI; 1992-150885/18.
XX P-PSDB; AAR23155.
XX
XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
XX activity - having conserved regions mutated or deleted, for use
XX in e.g. PCR, sequencing and detection assays
XX
XX Claim 11; Page 59; 185pp; English.
XX
XX The sequence is that of Thermus species sps17 polymerase DNA which
XX has been mutated. The mutation designated pSPSd2-151 causes the
XX polymerase enzyme produced to exhibit a different amt. of 5'-3'
XX exonuclease activity than the native enzyme. Thermostable DNA
XX polymerases are useful in many recombinant DNA techniques, esp.
XX nucleic acid amplification by PCR, self-sustained sequence
XX replication (SSR) and high temp. DNA sequencing. The absence of
XX 5'-3' nuclease activity may facilitate higher sensitivity allelic
XX discrimination in a combined polymerase ligase chain reaction (PLCR)
XX assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
XX in enzymes used in homogeneous assays for the amplification and
XX detection of a target nucleic acid sequence. Mutation of the DNA
XX encoding particular regions of the enzymes can be used to prepare
XX a range of recombinant proteins having 5'-3' exonuclease activity
XX to a complete lack of activity.
XX See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
XX SQ Sequence 2043 BP; 339 A; 697 C; 703 G; 304 T; 0 other;

Alignment Scores:
Pred. No.: 1.8 Length: 2043
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 13 Gaps: 0

US-09-823-649a-4 (1-11) x AAQ24008 (1-2043)
OY 1 LeuSer**GluSerIleProTyGluGlu 11
DB 1576 CTCTCCGAGAGCTCTCCATCCCTACGAGGAG 1608

RESULT 4
AAQ24007
ID AAQ24007 standard; DNA; 2277 BP.
XX
XX AAQ24007;
XX
XX 22-OCR-1992 (first entry)
XX
XX Mutant thermostable DNA polymerase pSPSd2-73.

```

```

XX 5'-3'; exonuclease; PCR, amplification; SSR, sequencing; PLCR; ss.
XX Thermus species sps17.
XX
XX Key Location/Qualifiers
XX old_sequence 1..2
FT /tag= a
FT /note= "nucleotides 4-219 deleted from the native
FT sequence."
XX
XX M09206200-A.
XX
XX 16-APR-1992.
XX
XX 30-SEP-1991; 91WO-US07035.
XX
XX 28-SEP-1990; 90US-0590213.
XX 28-SEP-1990; 90US-0590466.
XX 28-SEP-1990; 90US-0590490.
XX
XX (CETU ) CETUS CORP.
XX
XX Abramson RD, Gelfand DH;
XX
XX WPI; 1992-150885/18.
XX P-PSDB; AAR23154.
XX
XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
XX activity - having conserved regions mutated or deleted, for use
XX in e.g. PCR, sequencing and detection assays
XX
XX Claim 11; Page 59; 185pp; English.
XX
XX The sequence is that of Thermus species sps17 polymerase DNA which
XX has been mutated. The mutation designated pSPSd2-73 causes the
XX polymerase enzyme produced to exhibit a different amt. of 5'-3'
XX exonuclease activity than the native enzyme. Thermostable DNA
XX polymerases are useful in many recombinant DNA techniques, esp.
XX nucleic acid amplification by PCR, self-sustained sequence
XX replication (SSR) and high temp. DNA sequencing. The absence of
XX 5'-3' nuclease activity may facilitate higher sensitivity allelic
XX discrimination in a combined polymerase ligase chain reaction (PLCR)
XX assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
XX in enzymes used in homogeneous assays for the amplification and
XX detection of a target nucleic acid sequence. Mutation of the DNA
XX encoding particular regions of the enzymes can be used to prepare
XX a range of recombinant proteins having 5'-3' exonuclease activity
XX to a complete lack of activity.
XX See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
XX
XX SQ Sequence 2277 BP; 378 A; 779 G; 779 G; 341 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 2.04 Length: 2277
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 13 Gaps: 0

US-09-823-649a-4 (1-11) x AAQ24007 (1-2277)
OY 1 LeuSer**GluSerIleProTyGluGlu 11
DB 1810 CTCTCCGAGAGCTCTCCATCCCTACGAGGAG 1842

RESULT 5
AAQ24006
ID AAQ24006 standard; DNA; 2367 BP.
XX
XX AAQ24006;
XX
XX

```

FT sequence.
 XX WO9206200-A.
 XX 16-APR-1992.
 XX 30-SEP-1991; 91WO-US07035.
 XX 28-SEP-1990; 90US-0590213.
 XX 28-SEP-1990; 90US-0590466.
 XX 28-SEP-1990; 90US-0590490.
 XX (CETU) CETUS CORP.
 XX Abramson RD, Gelfand DH;
 XX WPI: 1992-150885/18.
 XX P-PSDB; AAR23157.
 XX
 XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 XX Claim 11; Page 59; 185pp; English.
 XX
 XX The sequence is that of Thermus species sps17 polymerase DNA which
 CC has been mutated. The mutation designated pSPSd288 causes the
 CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. Thermostable DNA
 CC polymerases are useful in many recombinant DNA techniques, esp.
 CC nucleic acid amplification by PCR, self-sustained sequence
 CC replication (SSR) and high temp. DNA sequencing. The absence of
 CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
 CC
 XX SQ Sequence 1635 BP; 267 A; 572 C; 559 G; 237 T; 0 other;
 XX
 XX Alignment Scores:
 Pred. No.: 1.39 length: 1635
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 96.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-823-649A-4 (1-11) x AAQ24010 (1-1635)
 QY 1 Leuser**GluSer**IleProTyrGluGln 11
 DB 1168 CTCTCCAGAGCTCTCCATCCCTACGAGGAG 1200
 RESULT 2
 AAQ24009
 ID AAQ24009 standard; DNA; 1899 BP.
 XX
 AC AAQ24009;
 XX
 XX 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase pSPSd2-199.
 XX
 XX 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
 OS Thermus species sps17.
 XX
 KW Key Location/Qualifiers
 XX

FT old_sequence 1..2
 FT /*tag= a
 FT /note= "nucleotides 4-597 deleted from the native
 FT sequence."
 XX
 XX WO9206200-A.
 XX 16-APR-1992.
 XX 30-SEP-1991; 91WO-US07035.
 XX 28-SEP-1990; 90US-0590213.
 XX 28-SEP-1990; 90US-0590466.
 XX 28-SEP-1990; 90US-0590490.
 XX (CETU) CETUS CORP.
 XX Abramson RD, Gelfand DH;
 XX WPI: 1992-150885/18.
 XX P-PSDB; AAR23156.
 XX
 XX Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 XX Claim 11; Page 59; 185pp; English.
 XX
 XX The sequence is that of Thermus species sps17 polymerase DNA which
 CC has been mutated. The mutation designated pSPSd2-199 causes the
 CC polymerase enzyme produced to exhibit a different amt. of 5'-3'
 CC exonuclease activity than the native enzyme. Thermostable DNA
 CC polymerases are useful in many recombinant DNA techniques, esp.
 CC nucleic acid amplification by PCR, self-sustained sequence
 CC replication (SSR) and high temp. DNA sequencing. The absence of
 CC 5'-3' nuclease activity may facilitate higher sensitivity allelic
 CC discrimination in a combined polymerase ligase chain reaction (PLCR)
 CC assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable
 CC in enzymes used in homogeneous assays for the amplification and
 CC detection of a target nucleic acid sequence. Mutation of the DNA
 CC encoding particular regions of the enzymes can be used to prepare
 CC a range of recombinant proteins having 5'-3' exonuclease activity
 CC to a complete lack of activity.
 CC See also AAQ23993-Q24013, AAQ24320-36 and AAQ24343-60.
 CC
 XX SQ Sequence 1899 BP; 320 A; 647 C; 647 G; 285 T; 0 other;
 XX
 XX Alignment Scores:
 Pred. No.: 1.65 length: 1899
 Score: 48.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 96.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-823-649A-4 (1-11) x AAQ24009 (1-1899)
 QY 1 Leuser**GluSer**IleProTyrGluGln 11
 DB 1432 CTCTCCAGAGCTCTCCATCCCTACGAGGAG 1464
 RESULT 3
 AAQ24008
 ID AAQ24008 standard; DNA; 2043 BP.
 XX
 AC AAQ24008;
 XX
 XX 22-OCT-1992 (first entry)
 XX
 DE Mutant thermostable DNA polymerase pSPSd2-151.
 XX
 XX 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR; ss.
 KW
 XX

```

FT      /product= DNA polymerase
PN      US5837450-A.
XX
XX      17-NOV-1998.
PD
XX
XX      06-JUN-1995; 95US-0471066.
PF
XX      06-JUN-1994; 94US-0254359.
PR      07-DEC-1992; 92US-0986330.
PR      04-JUN-1993; 93US-0073384.
PR      06-JUN-1995; 95US-0471066.
PA      (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
XX      Brow MAD, Dahlberg JE, Iyamichev VI;
PI      WPI: 1999-023438/02.
DR      P-PSDB: AAW80428.
XX
XX      Detection of target nucleic acid molecules - uses modified
PT      thermostable enzymes with specific cleavage activity to create
PT      specific detection products from oligo:nucleotide(s) and target
PT      hybridisation
XX
XX      Disclosure; Fig 2; 91pp; English.
PS
XX      The present sequence encodes a thermostable DNA polymerase enzyme.
CC      The sequence is used in the course of the invention. The specification
CC      describes a method for detecting the presence of a nucleic acid
CC      molecule. The method uses a thermostable 5' nuclease derived from a
CC      thermostable polymerase modified to have reduced synthetic activity,
CC      where the 5' nuclease is capable of cleaving a linear nucleic acid
CC      duplex structure to create a single, single-stranded cleavage product.
CC      The methods are used for the specific detection of nucleic acid
CC      sequences, via a cleavage-based procedure, but without the need for
CC      amplification of target sequences. Thermostable polymerases, altered to
CC      have nuclease, but not polymerase activity are preferably used due to
CC      their specificity. The cleavage product specifically formed is detected,
CC      preferably by the use of radioactively labelled oligonucleotides. These
CC      can be used in e.g. forensic testing or paternity determination.
XX
SQ      Sequence 2496 BP; 402 A; 830 C; 854 G; 410 T; 0 other;

Alignment Scores:
Pred. No.: 2.27      Length: 2496
Score: 48.00      Matches: 10
Percent Similarity: 90.91%      Conservative: 0
Best Local Similarity: 90.91%      Mismatches: 1
Query Match: 96.00%      Indels: 0
DB: 20      Gaps: 0

US-09-823-649A-4 (1-11) x AAV63401 (1-2496)
OY      1 Leuser**Glucose1leProTyrGluGlu 11
        ||||| ||||| ||||| ||||| |||||
DB      2029 CTCCTCCGGGAGCTTTCATCCCTACGAGCAG 2061

RESULT 10
AAQ92369
ID      AAQ92369 standard; DNA; 2512 BP.
XX
XX      AAQ92369;
AC
XX      28-OCT-1995 (first entry)
DT
XX      Tfil DNA-polymerase.
DE
XX      DNA-polymerase; Tfil; thermostable enzyme; ss.
XX      Thermus filiformis.
OS
XX      Key      Location/Qualifiers
FH

```

```

FT      CDS      9..2511
FT      /*tag= a
FT      /FC_number= 2.7.7.7
PN      WO9514770-A.
XX
XX      01-JUN-1995.
PD
XX
XX      23-NOV-1994; 94WO-N00135.
PF
XX      25-NOV-1993; 93WA-0250288.
PR      (PACI-) PACIFIC ENZYMES 1993 LTD.
PA
XX      Bergquist PI, Day DJ, Gibbs MD, Reeves RA, Saul DJ;
PI      WPI: 1995-206929/27.
DR      P-PSDB: AAR76060.
XX
XX      New heat-stable DNA polymerase from Thermus filiformis - has reverse
PT      transcriptase activity in the presence of magnesium ions.
PT      transcriptase activity in the presence of magnesium ions.
XX
XX      Claim 7; Fig 1; 40pp; English.
PS
XX      The sequence encodes thermostable DNA-polymerase Tfil (DNA
CC      nucleotidyltransferase DNA-directed enzyme) which has reverse-
CC      transcriptase activity in the presence of Mg2+. The enzyme is a
CC      versatile DNA-polymerase suitable for use in DNA synthesis from DNA
CC      and RNA templates, and for automated DNA sequencing. Both of these
CC      reactions are performed in the same buffer. The enzyme has a high
CC      salt tolerance and is also suitable for use in cycled fluorescent
CC      DNA sequencing using dye-primers and dye-terminators.
XX
SQ      Sequence 2512 BP; 415 A; 862 C; 854 G; 381 T; 0 other;

Alignment Scores:
Pred. No.: 2.29      Length: 2512
Score: 48.00      Matches: 10
Percent Similarity: 90.91%      Conservative: 0
Best Local Similarity: 90.91%      Mismatches: 1
Query Match: 96.00%      Indels: 0
DB: 16      Gaps: 0

US-09-823-649A-4 (1-11) x AAQ92369 (1-2512)
OY      1 Leuser**Glucose1leProTyrGluGlu 11
        ||||| ||||| ||||| ||||| |||||
DB      2034 CTCCTCCGGGAGCTTTCATCCCTACGAGCAG 2066

RESULT 11
AAQ86927
ID      AAQ86927 standard; DNA; 4947 BP.
XX
XX      AAQ86927;
AC
XX      05-DEC-1995 (first entry)
DT
XX      Tspis17 polymerase gene.
DE
XX      Polymerase; Thermus; enzyme; thermostable; exonuclease; ss.
XX      Thermus species sps17.
OS
XX      Key      Location/Qualifiers
FH      CDS      1246..3738
FT      /*tag= a
FT      /product= Thermostable polymerase.
XX
XX      US5405774-A.
XX      11-APR-1995.
PD
XX      22-AUG-1986; 86US-0899241.
PF

```



```

PR    10-SEP-1993;      93US-0119754.
PR    22-AUG-1986;     86US-0899241.
PR    17-JUN-1987;     87US-0063509.
PR    12-JAN-1988;     88US-0134411.
PR    15-MAY-1990;     90US-0523394.
PR    28-SEP-1990;     90US-0590213.
PA    (HOFF ) HOFFMANN LA ROCHE INC.
XX
XX    Abramson RD, Gelfand DH, Greenfield IL;
DR    WPI: 1995-154582/20.
DR    P-PSDB; AAR72875.
XX
PT    DNA for Thermus species sps17 DNA polymerase - modified to encode
XX    an N-terminal deletion to reduce 5'-3' exo-nuclease activity
XX
XX    Claim 3; Columns 5-12; 26pp; English.
XX
XX    This sequence encodes the thermostable polymerase of Thermus species
CC    sps17. It may be mutated to encode an Asp residue at position 43
CC    instead of the wild type Gly residue. This substitution is shown to
CC    result in a 100 fold decrease in the 5'-3' exonuclease activity with
XX    no change to the polymerase activity of the enzyme.
SQ    Sequence 4947 BP; 822 A; 1601 C; 1774 G; 750 T; 0 other;
Alignment Scores:
Pred. No.:          5           length:       4947
Score:              58.00        Matches:      10
Percent Similarity: 90.91%      Conservative: 10
Best Local Similarity: 90.91%   Mismatches:  1
Query Match:        96.00%      Indels:      0
DB:                  16         Gaps:         0
US-09-823-649A-4 (1-11) x AAQ86927 (1-4947)
OY    1 LeuSer***GluLeuSerIleProTyrGluGln 11
Db    3316 CTCCTCCAGAGGCTCTCCATCCCTCAGAGAG 3348
RESULT 12
AAQ80745
ID    AAQ80745 standard; DNA; 1600 Bp.
XX
XX    AAQ80745:
AC    19-JUL-1995 (first entry)
XX
DE    5' Nuclease from Taq DNAP.
XX
XX    DNA-polymerase; DNAP; Taq; DNA cleavage; RNA cleavage; 5' nuclease; ds
KM    Thermus aquaticus YT-1.
XX
OS    Thermus aquaticus YT-1.
XX
PN    MO9429482-A.
PD    22-DEC-1994.
PP    06-JUN-1994;      94MO-US06253.
PR    04-JUN-1993;      93US-0073384.
XX
XX    (THIR-) THIRD WAY TECHNOLOGIES INC.
XX
XX    Brow MAD, Dahlberg JE, Lyamichev VI;
PI    WE1: 1995-036504/05.
XX
XX    5' Nuclease(s) derived from thermostable DNA polymerase(s) - have
PT    cleavage activity with reduced synthetic ability, used for
PR    detection of specific target sequences.

```

```

XX Claim 5; Page 95-96; 159pp; English.
PS
CC The Tag DNA polymerase (DNAP) gene was amplified by PCR. Amplified
CC fragments were ligated into pTYO18 vector, which contains the
CC hybrid trp-lac (lac) promoter. An amplification/selection
CC protocol was used to isolate clone 4B containing a mutated Tag DNA
CC gene (mutTag) (sequence given in AA080746) having normal 5' nucleas
CC activity but less than 1% of the wt Tag DNAP activity. mutTag was
CC cut from pTYO18 by EcoRI-SalI digestion and cloned into pET-3c.
CC Thus clone was digested with BstXI and BamHI. The DNA was treated
CC with DNAPect Klenow fragment and dNTPs, BstI-ended and religated,
CC resulting in an in-frame deletion of 903 nucleotides. The DNA
CC sequence of the resulting 5' nucleas is given in AA080745.
CC
CC
CC
SQ Sequence 1600 BP; 287 A; 518 C; 553 G; 242 T; 0 other;
Alignment Scores:
Pred. No.: 5.97 length: 1600
Score: 45.00 Matches: 9
Percent Similarity: 90.01% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 16 Gaps: 0
US-09-823-649A-4 (1-11) x AA080745 (1-1600)
QY 1 Leuser**G1aLeuserFileProTyrglu11
||||| ||||||:|||||||
DB 1133 CTCGCCAGGACCTPAGCATCTTACGACGAG 1165
RESULT 13
AAT27686
ID AAT27686 standard; DNA; 1600 BP.
XX
XX AAT27686;
AC
XX
XX 18-NOV-1996 (first entry)
DT
XX
XX Mutant Thermus aquaticus DNA polymerase coding sequence (clone 4F)
DE
XX
XX p53; mutant; mutation; cleavage; nuclease; cleavage; Thermus;
KW Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
KW Staphylococcus; identification; detection; ds.
XX
XX Thermus aquaticus.
OS
XX
XX
XX Key location/qualifiers
FH CDS 14..1600
PT /atg a
FT /product= Mutant DNA polymerase.
XX
XX W09615267-A1.
PN
XX
XX 23-MAY-1996.
PD
XX
XX 09-NOV-1995; 95WO-US14673.
PE
XX
XX 30-AUG-1995; 95US-0520946.
PR 09-NOV-1994; 94US-0337184.
PR 09-MAR-1995; 95US-0402601.
PR 07-JUN-1995; 95US-0484956.
XX
XX (THIR-) THIRD WAVE TECHNOLOGIES INC.
PA
XX
XX Brow MAD, Dahlberg JE, Fors L, Heisler LM, Lyamichnev VI;
PI Oldenburg MC, Olive DM;
XX
XX WPI: 1996-259862/26.
DR
XX
XX P-PSDB: AAR96267.
PT
XX
XX Cleavage of nucleic acids to detect mutation(s) - allows detection
PT esp. in human p53 gene, to identify strains of microorganisms and

```

PT viruses

XX Example 2A; Page 257-258; 433pp; English.

XX
PS
CC Cleaveage of nucleic acids using an enzyme, especially a nuclease
CC selected from the group consisting of Cleavase (RTM) BN enzyme,
CC Thermus aquaticus DNA polymerase, and the Saccharomyces cerevisiae
CC polymerase, Escherichia coli ExoIII and the Thermophilus
CC Rad1/Rad10 complex. The nucleic acid substrate is preferably an
CC oligonucleotide containing a human p53 gene sequence or
CC alternatively, microbial gene sequences. Cleaveage products are
CC compared to the cleavage products of reference gene sequences. The
CC method is used for detecting mutation in the human p53 gene; for
CC identifying strains of microorganisms, especially bacteria selected
CC from the group of members of the genera Campylobacter,
CC Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
CC The method may also be used for the identification of viruses,
CC especially hepatitis C virus and simian immunodeficiency virus.
CC Thermus aquaticus (Taq) DNA polymerase was amplified using two
CC primers (AAT70347, AAT70348). The Taq polymerase DNA was inserted into
CC the BamHI restriction site of the expression vector pET-3c and mutant
CC genes were created from that construct. This mutant was created
CC after the vector was digested with BstXI and BamHI. The DNA was
CC then treated with the Klenow fragment of DNAPol to trim both 3'
CC overhangs to blunt ends which were then ligated together, resulting
CC in an in-frame deletion of 903 nucleotides. The resulting mutant
CC Taq polymerase is also referred to as the Cleavase BB enzyme.

XX Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;

Alignment Scores:

Pred. No.: 5.97 Length: 1600
Score: 45.00* Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 17 Gaps: 0

US-09-823-649A (1-11) x AAT70347 (1-1600)

QY 1 LeuSer**GluLeuSerIleProTyrGluGlu 11

DB 1133 CTCCTCCAGAGCTACGATCCCTTACGAGGAG 1165

RESULT 14

AAT70347 ID AAT70347 standard; DNA: 1600 BP.

XX AAT70347;

XX 03-APR-1998 (first entry)

XX Synthesis deficient Taq DNA polymerase gene clone 4F.

XX DNA polymerase: taq polymerase gene; DMP: 5' nuclease activity;

XX Cleavase BB; DNA cleavage; reduced synthetic activity; ss.

XX Synthetic.

XX Thermus aquaticus.

XX US5614402-A.

XX 25-MAR-1997.

XX 06-JUN-1994; 94US-0254359.

XX 07-DEC-1992; 92US-0986330.

XX 04-JUN-1993; 93US-0073384.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Brow MAD, Dahlberg JE, Lyamichev VI;

XX

DR WPI; 1997-201481/18.

XX Thermostable 5' nuclease derived from thermostable polymerase - has

XX reduced synthetic activity useful in nucleic acid detection assays

XX Example 2A; Columns 79-82; 93pp; English.

XX
PS
CC The present sequence, clone 4F, encodes a Taq DNA polymerase enzyme
CC that has normal 5' nuclease activity, but reduced synthetic ability.
CC This sequence was derived from the multigene construct described in
CC AAT70343. The entire multigene was cut from the plasmid, and cloned
CC into pET-3c. This clone was digested with BstXI and BamHI, at unique
CC sites. The 3' overhang of BstXI was trimmed to a blunt end, while the 5'
CC overhang of BamHI was filled in. The blunt ends were ligated together.
CC This resulted in an in-frame deletion of 903 nucleotides. The enzyme
CC encoded by the present protein is also referred to as Cleavase BB.

XX Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;

Alignment Scores:

Pred. No.: 5.97 Length: 1600
Score: 45.00* Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 90.00% Indels: 0
DB: 18 Gaps: 0

US-09-823-649A-4 (1-11) x AAT70347 (1-1600)

QY 1 LeuSer**GluLeuSerIleProTyrGluGlu 11

DB 1133 CTCCTCCAGAGCTACGATCCCTTACGAGGAG 1165

RESULT 15

AAT70347 ID AAT70347 standard; DNA: 1600 BP.

XX AAT70347;

XX 14-APR-1998 (first entry)

XX Taq gene 5' nuclease clone 3F (Cleavase BB).

XX Nucleic acid cleavage; DNA cleavage; RNA cleavage; 5' nuclease;

XX Taq; DNA polymerase; Cleavase BB; ds.

XX Thermus aquaticus YT-1.

XX Synthetic.

XX W09/27214-A1.

XX 31-JUL-1997.

XX 22-JAN-1997; 97MO-US01072.

XX 02-DEC-1996; 96US-0758038.

XX 24-JAN-1996; 96US-0599491.

XX 12-JUL-1996; 96US-0682853.

XX 29-NOV-1996; 96US-0756386.

XX 02-DEC-1996; 96US-0758314.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Brow MAD, Dahlberg JE, Hall JG, Kaiser KM, Lyamichev VI;

XX Olive DM, Prudent DR;

XX WPI; 1997-393613/36.

XX Thermostable structure-specific nucleases(s) - used for detection and

XX characterisation of nucleic acid sequences and variations in nucleic

XX acid sequences

XX Example 2; Page 245; 457pp; English.

XX This DNA sequence, denoted clone 3f or Cleavase BB, comprises an
 CC altered *Thermus aquaticus* DNA polymerase (Taq) gene in which
 CC nucleotides 875-1778 of the wild-type gene coding sequence are
 CC deleted. Mutant gene mutTaq (see AAT76643) was used as the starting
 CC material for the construct. Cleavase BB is a thermostable
 CC structure-specific nuclease preferred for use in nucleic acid
 CC cleavage methods of the invention. Mutant genes (AAT76644-47)
 CC were constructed in order to determine which portions of the Taq
 CC polymerase domain can be altered without eliminating 5' nuclease
 CC activity. The invention relates to means for the detection and
 CC characterisation of nucleic acid (NA) sequences and variations in
 CC NA. It relates to methods for forming a NA cleavage structure on
 CC a target sequence and cleaving the NA cleavage structure in a
 CC site-specific manner. The 5' nuclease activity of various enzymes
 CC (see AAM24210-13) is used to cleave the target-dependent cleavage
 CC structure, thereby indicating the presence of specific NA sequences
 CC or specific variations of them.

XX
 SQ Sequence 1600 BP; 286 A; 519 C; 553 G; 242 T; 0 other;

Alignment Scores:

Pred. No.:	5.97	Length:	1600
Score:	45.00	Matches:	9
Percent Similarity:	90.91%	Conservative:	1
Best Local Similarity:	81.82%	Mismatches:	1
Query Match:	90.00%	Indels:	0
DB:	18	Gaps:	0

US-09-823-649A-4 (1-11) x AAT76647 (1-1600)

QY 1 LeuSer**GIuLeuSerIleProTyrGIuIu 11
 DB 1133 CTCTCCAGAGACTAGCCATCCCTTACGAGAG 1165

Search completed: January 15, 2003, 12:55:32
 Job time : 134.143 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 12:36:21 ; Search time 26 Seconds
(without alignments)
129.748 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSKELSTPYEE 11

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 44362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -
-MODEL=framer+g2n.model -DEV=xlh
-O=/cg2_1/USPTO_spool/US09823649/runat_14012003_151002_29127/app_query.fasta.1.1393
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPL=0
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649/@CGN.1.1_35_@runat_14012003_151002_29127 -MCPD=6 -ICPD=3
-NO.XIPXY -NO.MMAP -LARGEORDERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMOUT=120
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents NA.*
1: /cg2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cg2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cg2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cg2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cg2_6/ptodata/1/ina/PCRTUS.COMB.seq.*
6: /cg2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	48	96.0	2493	1 US-07-977-434-5	Sequence 5, Appl 1
2	48	96.0	2493	1 US-08-458-819-5	Sequence 5, Appl 1
3	48	96.0	2493	1 PCT-US91-07035-5	Sequence 5, Appl 1
4	48	96.0	2496	1 US-08-073-384C-2	Sequence 2, Appl 1
5	48	96.0	2496	1 US-08-254-359A-2	Sequence 2, Appl 1
6	48	96.0	2496	1 US-08-483-043-2	Sequence 2, Appl 1
7	48	96.0	2496	1 US-08-481-238-2	Sequence 2, Appl 1
8	48	96.0	2496	1 US-08-471-066B-2	Sequence 2, Appl 1
9	48	96.0	2496	2 US-08-484-956-2	Sequence 2, Appl 1
10	48	96.0	2496	2 US-08-757-653-2	Sequence 2, Appl 1
11	48	96.0	2496	2 US-08-599-491-2	Sequence 2, Appl 1
12	48	96.0	2496	2 US-08-756-386-2	Sequence 2, Appl 1

13	48	96.0	2496	2 US-08-823-516-2	Sequence 2, Appl 1
14	48	96.0	2496	3 US-08-682-853A-2	Sequence 2, Appl 1
15	48	96.0	2496	3 US-08-759-038-2	Sequence 2, Appl 1
16	48	96.0	2496	3 US-08-758-314-2	Sequence 2, Appl 1
17	48	96.0	2496	4 US-09-350-309-2	Sequence 2, Appl 1
18	48	96.0	2496	4 US-08-520-946-2	Sequence 2, Appl 1
19	48	96.0	1600	1 US-08-073-384C-12	Sequence 12, Appl 1
20	45	90.0	1600	1 US-08-254-359A-12	Sequence 12, Appl 1
21	45	90.0	1600	1 US-08-483-043-12	Sequence 12, Appl 1
22	45	90.0	1600	1 US-08-471-066B-12	Sequence 12, Appl 1
23	45	90.0	1600	2 US-08-484-956-12	Sequence 12, Appl 1
24	45	90.0	1600	2 US-08-757-653-12	Sequence 12, Appl 1
25	45	90.0	1600	2 US-08-599-491-12	Sequence 12, Appl 1
26	45	90.0	1600	2 US-08-756-386-12	Sequence 12, Appl 1
27	45	90.0	1600	2 US-08-823-516-12	Sequence 12, Appl 1
28	45	90.0	1600	3 US-08-682-853A-12	Sequence 12, Appl 1
29	45	90.0	1600	3 US-08-759-038-12	Sequence 12, Appl 1
30	45	90.0	1600	3 US-08-758-314-12	Sequence 12, Appl 1
31	45	90.0	1600	4 US-09-350-309-12	Sequence 12, Appl 1
32	45	90.0	1600	4 US-08-520-946-12	Sequence 12, Appl 1
33	45	90.0	1600	4 US-09-096-399-1	Sequence 1, Appl 1
34	45	90.0	1682	3 US-08-648-657-1	Sequence 1, Appl 1
35	45	90.0	1686	2 US-08-648-657-14	Sequence 14, Appl 1
36	45	90.0	1686	2 US-08-648-657-12	Sequence 2, Appl 1
37	45	90.0	1689	2 US-08-648-657-2	Sequence 2, Appl 1
38	45	90.0	1696	3 PCT-US95-15327-3	Sequence 3, Appl 1
39	45	90.0	1794	5 PCT-US95-15327-3	Sequence 3, Appl 1
40	45	90.0	2499	1 US-07-977-434-1	Sequence 1, Appl 1
41	45	90.0	2499	1 US-08-458-819-1	Sequence 1, Appl 1
42	45	90.0	2499	3 US-07-602-848E-1	Sequence 1, Appl 1
43	45	90.0	2499	3 US-09-587-856-1	Sequence 1, Appl 1
44	45	90.0	2499	4 US-09-777-537-1	Sequence 1, Appl 1
45	45	90.0	2499	4 US-09-777-537-1	Sequence 1, Appl 1

ALIGNMENTS

RESULT 1
US-07-977-434-5
; Sequence 5, Application US/07977434
; Patent No. 5466591
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David H.
; APPLICANT: Abramson, Richard D.
; TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: 7
; SOFTWARE: wordperfect 2.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,434
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,490
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,466
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,213
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 523,394

FILED DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseert
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 5466591 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORGANISM: Thermus species sps17
FEATURE
NAME/KEY: CDS
LOCATION: 1..2490
US-07-977-434-5
US-07-977-434-5

Alignment Scores:
Pred. No.: 0.566 Length: 2493
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0

US-09-823-649A-4 (1-11) x US-07-977-434-5 (1-2493)
Cy 1 leuser**GluLeuSerIleProTyrGluGlu 11
||||| ||||||| ||||||| ||||||| |||||||
Db 2026 CTTCTCCAGAGACTCTCATCCCTTCCAGAGAG 2058

RESULT 2
US-08-458-819-5
Sequence 5, Application US/08458819
Patent No. 5795762
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES

NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: wordperfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,819
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/977,434
FILING DATE: 23-FEB-1993
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cseert
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 5795762 8753
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:

ORGANISM: Thermus species sps17
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2490
US-08-458-819-5
Alignment Scores:
Pred. No.: 0.566 Length: 2493
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
US-09-823-649a-4 (1-11) x US-08-458-819-5 (1-2493)
QY 1 LeuSer**GluLeuSerIleProTyrGlnGlu 11
Db 2026 CTCCTCCAGAGACTCTCATCCCTAGAGAGAG 2058
RESULT 3
PCT-US91-07035-5
Sequence 5, Application PC/US910/035
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Adams, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cetus Corporation
STREET: 1400 Fifty-third street
CITY: Emeryville
STATE: California
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/07035
FILING DATE: 19910930
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US90/07641
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471

FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Sias Ph.D., Stacey R.
REGISTRATION NUMBER: 32,630
REFERENCE/DOCKET NUMBER: Case No. 2580
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-420-3300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2493 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Thermus species sps17
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2490
PCT-US91-07035-5
Alignment Scores:
Pred. No.: 0.566 Length: 2493
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
US-09-823-649a-4 (1-11) x PCT-US91-07035-5 (1-2493)
QY 1 LeuSer**GluLeuSerIleProTyrGlnGlu 11
Db 2026 CTCCTCCAGAGACTCTCATCCCTAGAGAGAG 2058
RESULT 4
US-08-073-384C-2
Sequence 2, Application US/0807384C
Patent No. 5541411
GENERAL INFORMATION:
APPLICANT: Dahlberg, James R.
APPLICANT: Lyamichov, Victor I.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
TITLE OF INVENTION: POLYMERASE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVKSTOCK, MEDLEN & CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,384C
FILING DATE: 04-JUN-1993
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-00613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-073-384C-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-4 (1-11) x US-08-073-384C-2 (1-2496)
QY 1 Leuser***GlueSerIleProTyrGluGlu 11
||||| ||||||||| |||||||||
Db 2029 CTCGCCGGGAGCTTCCATCCCTACGACGAG 2061

RESULT 5
US-08-254-359A-2
Sequence 2, Application US/08254359A
Patent No. 5614402
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: BROW, MARY ANN D.
TITLE OF INVENTION: 5' NUCLEASES DERIVED FROM THERMOSTABLE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSER: HAVERSHOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,359A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-254-359A-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-4 (1-11) x US-08-254-359A-2 (1-2496)
QY 1 Leuser***GlueSerIleProTyrGluGlu 11
||||| ||||||||| |||||||||
Db 2029 CTCGCCGGGAGCTTCCATCCCTACGACGAG 2061

RESULT 6
US-08-483-043-2
Sequence 2, Application US/08483043
Patent No. 5691142
GENERAL INFORMATION:
APPLICANT: DAHLBERG, JAMES E.
APPLICANT: BROW, MARY ANN D.
TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: HAVERSHOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,043
FILING DATE: 06-JUN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-00613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-483-043-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496

Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-4 (1-11) x US-08-483-043-2 (1-2496)

QY 1 Leuser***GlueSer1leProTyrglu 11
||||| ||||||| ||||||| |||||||

Db 2029 CTCCTCGGAGGAGCTTCCATCCCTACGAGAG 2061

RESULT 7

US-08-481-238-2

; Sequence 2, Application US/08481238

; Patent No. 5795763

; GENERAL INFORMATION:

; APPLICANT: DAHLBERG, JAMES E.

; APPLICANT: LYAMICHEV, VICTOR I.

; APPLICANT: BROW, MARY ANN D.

; TITLE OF INVENTION: SYNTHESIS-DEFICIENT THERMOSTABLE DNA

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MEDLEN & CARROLL

; STREET: 220 MONTGOMERY STREET, SUITE 2200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/481,238

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: CARROLL, PETER G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: F08S-01798

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2496 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-481-238-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 1 Gaps: 0

US-09-823-649A-4 (1-11) x US-08-481-238-2 (1-2496)

QY 1 Leuser***GlueSer1leProTyrglu 11
||||| ||||||| ||||||| |||||||

Db 2029 CTCCTCGGAGGAGCTTCCATCCCTACGAGAG 2061

RESULT 8

US-08-471-066B-2

; Sequence 2, Application US/08471066B

; Patent No. 5837450

GENERAL INFORMATION:
; APPLICANT: Dahlberg, James E.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: 5' Nucleases Derived From Thermostable
; TITLE OF INVENTION: DNA Polymerase
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,066B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/254,359

FILING DATE: 06-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/073,384

FILING DATE: 04-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/986,330

FILING DATE: 07-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: F08S-01800

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2496 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 2 Gaps: 0

US-09-823-649A-4 (1-11) x US-08-471-066B-2 (1-2496)

QY 1 Leuser***GlueSer1leProTyrglu 11
||||| ||||||| ||||||| |||||||

Db 2029 CTCCTCGGAGGAGCTTCCATCCCTACGAGAG 2061

RESULT 9

US-08-484-956-2

; Sequence 2, Application US/08484956

; Patent No. 5843654

; GENERAL INFORMATION:

; APPLICANT: DAHLBERG, JAMES E.

; APPLICANT: LYAMICHEV, VICTOR I.

; APPLICANT: BROW, MARY ANN D.

; APPLICANT: OLDENBURG, MARY C.

; APPLICANT: HRIJSER, LAURA

; TITLE OF INVENTION: DETECTION OF P53 MUTATIONS

NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAVERSTOCK, MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,956
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,601
FILING DATE: 09-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,164
FILING DATE: 09-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/254,359
FILING DATE: 06-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/073,384
FILING DATE: 04-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,330
FILING DATE: 07-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, J, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: FORS-01801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-956-2
Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 2
US-09-823-649a-4 (1-11) x US-08-484-956-2 (1-2496)
QY 1 Leuser**GluLeuserIleProTyGluIu 11
||||| |||||||||
Db 2029 CTTCCGGGAGCTTCCATCCCTACGAGAG 2061
RESULT 10
US-08-757-653-2
Sequence 2, Application US/08/57653
Patent No. 5843669
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,653
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02565
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-653-2
Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 2
US-09-823-649a-4 (1-11) x US-08-757-653-2 (1-2496)
QY 1 Leuser**GluLeuserIleProTyGluIu 11
||||| |||||||||
Db 2029 CTTCCGGGAGCTTCCATCCCTACGAGAG 2061
RESULT 11
US-08-599-491-2
Sequence 2, Application US/08/59491
Patent No. 5846717
GENERAL INFORMATION:
APPLICANT: BROW, MARY ANN D.
APPLICANT: GROTHUESCHEN HALL, JEFF S.
APPLICANT: LYAMICHEV, VICTOR
APPLICANT: OLIVE, DAVID W.
APPLICANT: PRUDENT, JAMES R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
TITLE OF INVENTION: INVADER-DIRECTED CLEAVAGE
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,491

FILED DATE: 23-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: P-40,027
REFERENCE/DOCKET NUMBER: FORS-01802
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-599-491-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 2

US-09-823-649A-4 (1-11) x US-08-599-491-2 (1-2496)
QY 1 Leuser**GluLeuserIleProTyrgIuGlu 11
||||| |||||||||
Db 2029 CTCGCGGAGAGCTTCCATCCCTACGAGAG 2061

RESULT 12
US-08-756-386-2
Sequence 2, Application US/08756386
Patent No. 5985557
GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,386
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02564
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-756-386-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 2

US-09-823-649A-4 (1-11) x US-08-756-386-2 (1-2496)
QY 1 Leuser**GluLeuserIleProTyrgIuGlu 11
||||| |||||||||
Db 2029 CTCGCGGAGAGCTTCCATCCCTACGAGAG 2061

RESULT 13
US-08-823-516-2
Sequence 2, Application US/08823516
Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PC/US97/01072
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-823-516-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 2

US-09-823-649a-4 (1-11) x US-08-823-516-2 (1-2496)

QY 1 Leuser**Glueuser1leProtyrCluClu 11
||||| ||||||| ||||||| ||||||| |||||||
Db 2029 CTCCTCCGGGAGCTTCCATCCCTACGAGG 2061

RESULT 14
US-08-682-853A-2
Sequence 2, Application US/08682853A
Patent No. 6001567

GENERAL INFORMATION:
APPLICANT: Brow, Mary Ann D.
APPLICANT: Grotelueschen Hall, Jeff S.
APPLICANT: Lyamichev, Victor
APPLICANT: Olive, David W.
APPLICANT: Prudent, James R.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY
TITLE OF INVENTION: INVADER-DIRECTED CLEAVAGE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/682,853A
FILING DATE: 12-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02306
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-682-853A-2

Alignment Scores:

Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 3

US-09-823-649a-4 (1-11) x US-08-682-853A-2 (1-2496)

QY 1 Leuser**Glueuser1leProtyrCluClu 11
||||| ||||||| ||||||| ||||||| |||||||
Db 2029 CTCCTCCGGGAGCTTCCATCCCTACGAGG 2061

RESULT 15
US-08-759-038-2
Sequence 2, Application US/08759038
Patent No. 6090543

GENERAL INFORMATION:
APPLICANT: Prudent, James R.
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Brow, Mary Ann D.
APPLICANT: Dahlberg, James E.
TITLE OF INVENTION: Cleavage Of Nucleic Acids
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,038
FILING DATE: 02-DEC-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02574
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2496 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-759-038-2

Alignment Scores:
Pred. No.: 0.567 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
Gaps: 0
DB: 3

US-09-823-649A-4 (1-11) x US-08-759-038-2 (1-2496)

Oy 1 LeuSer**GluLeuSerIleProTyrGluGlu 11
||||| ||||||||||||||||
Db 2029 CTCCTCGGAGACTTCCATCCCTACGAGAG 2061

Search completed: January 15, 2003, 12:58:46
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 Seconds

(without alignments)
72,584 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSTPYEE 11

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	96.0	831	1	DPOF_THETH
2	45	90.0	832	1	DPOI_THETH
3	45	90.0	834	1	DPOI_THETH
4	45	90.0	834	1	DPOI_THETH
5	39	78.0	229	1	CG18_YEAST
6	37	74.0	499	1	C1W5_HUMAN
7	37	74.0	1136	1	P01G_OMY
8	36	72.0	149	1	VE6_HPV31
9	36	72.0	327	1	TH14_ASPOR
10	36	72.0	753	1	NEC1_MOUSE
11	35	70.0	238	1	RM08_YEAST
12	35	70.0	752	1	NEC1_RAT
13	35	70.0	956	1	DPOI_THETH
14	34	68.0	326	1	TH14_YEAST
15	34	68.0	433	1	Y103_MERJA
16	34	68.0	859	1	RPA2_MERJA
17	33	66.0	162	1	PUR6_BACSU
18	33	66.0	197	1	Y828_MERJA
19	33	66.0	227	1	ID11_HUMAN
20	33	66.0	227	1	ID11_MESAU
21	33	66.0	227	1	ID11_MOUSE
22	33	66.0	227	1	ID11_RAT
23	33	66.0	403	1	ACKA_HELPJ
24	33	66.0	415	1	TRK3_MOUSE
25	33	66.0	753	1	NEC1_HUMAN
26	33	66.0	794	1	CADC_HUMAN
27	32	64.0	144	1	VP66_BPA18
28	32	64.0	181	1	TPPC_HALN1
29	32	64.0	208	1	Y101_ECOLI
30	32	64.0	262	1	YD5_CABEL
31	32	64.0	380	1	DRR_LISIN
32	32	64.0	390	1	MOEA_SYVA3
33	32	64.0	438	1	VATB_CHLMO

34	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
35	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
36	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
37	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
38	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
39	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
40	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
41	32	64.0	438	1	VATB_CHLMO	Q97997 Chlamydia p
42	31	62.0	215	1	RS3A_THETH	Q97822 Thermoplasma
43	31	62.0	222	1	KCV_AQUAE	Q67907 Aquifex aeo
44	31	62.0	266	1	AROE_HELPJ	Q97388 Helicobacter
45	31	62.0	274	1	TH14_AERPE	Q97920 Aeropyrum p

ALIGNMENTS

RESULT 1
DPOF_THETH ID DPOF_THETH STANDARD: PRT: 831 AA.
AC P30313:
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (Tfi polymerase 1).
GN POLA OR POL.
OS Thermus thermophilus.
OC Bacteria: Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ACM B-1257;
RX MEDLINE=93087201; PubMed=1454544;
RA Akhmetzhanov A.A., Vakhitov V.A.;
RT "Molecular cloning and nucleotide sequence of the DNA polymerase gene
from Thermus flavus";
RL Nucleic Acids Res. 20:5839-5839(1992).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA](n).
CC -1- BIOTECHNOLOGY: Used in the PCR method because of its high
thermostability.
CC -1- MISCELLANEOUS: Has an optimal temperature above 70 degrees
Celsius; can be used at temperatures up to 95 degree Celsius.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: X66105; CAA46900.1; -
DR PIR: S24929; S24929.
DR PIR: S26675; S26675.
DR HSP: P19821; 1190.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N1.
DR InterPro: IPR003583; HHH_1.
DR InterPro: IPR003584; HHH_2.
DR InterPro: IPR001532; XPGC_Rad.
DR Pfam: PF00476; DNA_Pol_A; 1.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF02739; 5_3_exonuclease; 1.
DR PRINTS: PR00868; DNA_POL.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; HHH1; 1.
DR SMART: SM00279; HHH2; 1.
DR SMART: SM00482; POLAC; 1.

DR SMART: SM00485; XPGN: 1.
DR TIGRFSM: TIGR00593; POLA: 1.
DR PROSITE: PS00447; DNA POLYMERASE_A: 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding.
FT DOMAIN 409 831 POLYMERASE (BY SIMILARITY).
SQ SEQUENCE 831 AA; 93783 MW; 96F93CFFA3CA530D CRC64;

Query Match 96.0%; Score 48; DB 1; Length 831;
Best Local Similarity 90.9%; Pred. No. 0.074;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSEXSLIPYEE 11
Db 677 LSEXSLIPYEE 687

RESULT 2
DPOL_THEAQ STANDARD; PRT; 832 AA.
ID DPOL_THEAQ
AC P19821;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase I, thermostable (EC 2.7.7.7) (Tag polymerase 1).
GN POLA OR POLI.
OS Thermus aquaticus.
OC Bacteria; Thermus/Delnococcus group; Delnococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=271;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=69197950; PubMed=2649500;
RA Lawyer F.C., Stoffel S., Saiki R.K., Myambo K., Drummond R.,
RA Gelfand D.H.;
RT "Isolation, characterization, and expression in Escherichia coli of
the DNA polymerase gene from Thermus aquaticus.";
RL J. Biol. Chem. 264:6427-6437(1989).
RN 12
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-YT1;
RX MEDLINE=95204371; PubMed=7896728;
RA Ishino Y., Ueno T., Miyaqi M., Uemori T., Imanura M., Tsunawake S.,
RA Kato I.;
RT "Overproduction of Thermus aquaticus DNA polymerase and its
structural analysis by ion-spray mass spectrometry.";
RL J. Biochem. 116:1019-1024(1994).
RN 13
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=95364959; PubMed=7637814;
RA Kim Y., Eom S.H., Wang J., Lee D.-S., Suh S.W., Stellz T.A.;
RT "Crystal structure of Thermus aquaticus DNA polymerase.";
RL Nature 376:612-616(1995).
RN 14
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 290-832.
RX MEDLINE=96016150; PubMed=7568114;
RA Korolev S., Nayal M., Barnes W.M., di Cera E., Waksman G.;
RT "Crystal structure of the large fragment of Thermus aquaticus DNA
polymerase I at 2.5-A resolution: structural basis for
thermostability.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9264-9268(1995).
RN 15
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX MEDLINE=96353962; PubMed=8717047;
RA Eom S.H., Wang J., Stellz T.A.;
RT "Structure of Tag polymerase with DNA at the polymerase active site.";
RL Nature 382:278-281(1996).
RN 16
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 295-832.
RX MEDLINE=99077817; PubMed=9857206;
RA Li Y., Korolev S., Waksman G.;
RT "Crystal structures of open and closed forms of binary and ternary
complexes of the large fragment of Thermus aquaticus DNA polymerase

RT I: structural basis for nucleotide incorporation.";
RN EMBJ J. 17:7514-7525(1998).
RN 17
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 290-832.
RX MEDLINE=98266352; PubMed=9605316;
RA Li Y., Kong Y., Korolev S., Waksman G.;
RT "Crystal structures of the Klenow fragment of Thermus aquaticus DNA
polymerase I complexed with deoxyribonucleoside triphosphates.";
RN Proc. Natl. Acad. Sci. U.S.A. 95:1116-1123(1998).
RN 18
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=98445410; PubMed=9770525;
RA Marali R., Sharkey D.J., Pais J.L., Murthy H.M.K.;
RT "Crystal structure of Tag DNA polymerase in complex with an inhibitory
Tag: the Fab is directed against an intermediate in the helix-coil
dynamics of the enzyme.";
RN Proc. Natl. Acad. Sci. U.S.A. 95:12562-12567(1998).
RN 19
RP X-RAY CRYSTALLOGRAPHY (2.30 ANGSTROMS) OF 293-831.
RX MEDLINE=99380545; PubMed=10449720;
RA Li Y., Mitaxov Y., Waksman G.;
RT "Structure-based design of Tag DNA polymerases with improved
properties of dideoxynucleotide incorporation.";
RN Proc. Natl. Acad. Sci. U.S.A. 96:9491-9496(1999).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ (DNA)(n).
CC -1- BIOTECHNOLOGY: used in the PCR method because of its high
thermostability. Has a relatively high error rate because it lacks
exonuclease proofreading functionality.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL: J04639; AAA27507.1; -
DR EMBL: D32013; BAA06775.1; -
DR PIR: A33530; A33530.
DR PDB: 1TAQ; 07-DEC-96.
DR PDB: 1TAU; 11-APR-97.
DR PDB: 1KTU; 08-NOV-96.
DR PDB: 2KTO; 22-DEC-99.
DR PDB: 3KTO; 22-DEC-99.
DR PDB: 4KTO; 22-DEC-99.
DR PDB: 5KTO; 30-SEP-98.
DR PDB: 1HGX; 14-OCT-98.
DR PDB: 1LXE; 12-MAR-97.
DR PDB: 1QSS; 16-AUG-99.
DR PDB: 1QGY; 16-AUG-99.
DR PDB: 1QFM; 16-AUG-99.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR001098; DNA_pol.
DR InterPro: IPR002298; DNA_pol.
DR InterPro: IPR000513; Exo_N.I.
DR InterPro: IPR003583; HNH_1.
DR InterPro: IPR003584; HNH_2.
DR Pfam: PF00476; DNA_pol_A_1.
DR Pfam: PF01367; 5_3_exonuclease_1.
DR Pfam: PF02739; 5_3_exonuclease_1.
DR PRINTS: PRO0668; DNAPOLI.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00278; HNH1; 1.
DR SMART: SM00279; HNH2; 1.
DR SMART: SM00482; POLA; 1.
DR TIGRFSM: TIGR00593; POLA: 1.
DR PROSITE: PS00447; DNA POLYMERASE_A: 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding; 3D-structure.
FT DOMAIN 410 832 POLYMERASE (BY SIMILARITY).

DB 680 LSOELAIPEE 690

RESULT 5

CG18_YEAST STANDARD; PRT: 229 AA.

AC P38794;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE G1/S-specific cyclin PCU5.

GN PCU5 OR YHR071W

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID:4932;

RN [1]

RP SOURCE FROM N.A.

RC STRAIN:8288C / AB972;

RC MEDLINE:94378003; PubMed:8091229;

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,

RA Du Z., Favell A., Fulton L., Gattung S., Gelsel C., Kistlen J.,

RA Kacaba T., Hillier L., Jier M., Johnston L., Langston Y.,

RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,

RA Nham M., Rifkin L., Riles L., St Peter H., Trevasaki E., Vaughan K.,

RA Vignati D., Wilcox L., Woldman P., Waterston R., Wilson R.,

RA Vaudin M.

RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome

VI11."

RL Science 265:2077-2082(1994).

CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. MOST SIMILAR TO G1/S-

SPECIFIC CYCLINS PCU4 AND PCU2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U00061; AAB68375.1;

DR PTR: S46696; S46696.

DR SGD: S000113; PCU5.

DR InterPro: IPR004366; Cyclin.

DR Pfam: PF00134; cyclin.1.

DR SMART: SMO0385; CYCLIN.1.

DR PROSITE: PS00292; CYCLINS; FALSE NEG.

DR Cyclin: Cell cycle; Cell division.

SO SEQUENCE 229 AA; 26467 MW; B6839AB9DB5D049 CRC64;

Query Match 78.0%; Score 39; DB 1; Length 229;

Best Local Similarity 63.6%; Pred. No. 1.2;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSELSIPEE 11

DB 174 LNEELAIPEE 184

RESULT 6

CIW5_HUMAN

ID CIW5_HUMAN

AC O95279;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Potassium channel subfamily K member 5 (Acid-sensitive potassium

channel protein TASK-2) (TWIK-related acid-sensitive K+ channel 2).

GN KCKS OR TASK2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Kidney;

RC MEDLINE:99030343; PubMed:9812978;

RA Reyes R., Duprat F., Desage F., Pink M., Salinas M., Harnan N.,

RA Lazdunski M.;

RT "Cloning and expression of a novel pH-sensitive two pore domain K+

channel from human kidney."

RL J. Biol. Chem. 273:30863-30869(1998).

CC -1- FUNCTION: PH DEPENDENT, VOLTAGE INSENSITIVE, OUTWARDLY RECTIFYING

POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS LOST AT HIGH EXTERNAL

K+ CONCENTRATIONS.

CC -1- SUBUNIT: HOMODIMER (POTENTIAL).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC IN LIVER, PLACENTA AND SMALL INTESTINE. IN THE KIDNEY, EXPRESSION

IS RESTRICTED TO THE DISTAL TUBULES AND COLLECTING DUCTS. NOT

EXPRESSED IN PROXIMAL TUBULES OR DUCTS.

CC -1- MISCELLANEOUS: INHIBITED BY QUININE, QUININDINE AND EXTERNAL

ACIDIFICATION.

CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM

CHANNELS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF084830; AAC79458.1;

DR Genew: HGNC:6280; KCKS5.

DR InterPro: IPR003280; Kichannel_2pore.

DR InterPro: IPR001622; Kichannel_pore.

DR InterPro: IPR000636; M-channel_1l9.

DR Pfam: PF00520; Ion_trans.1.

DR PRINTS: PR01333; ZPORECHANNEL.

KW Ionic channel; Transmembrane; Ion transport; Potassium transport;

KW Glycoprotein.

FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 8 26 POTENTIAL.

FT DOMAIN 85 112 PORE-FORMING 1 (POTENTIAL).

FT TRANSSEM 113 133 POTENTIAL.

FT DOMAIN 134 157 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 158 180 POTENTIAL.

FT DOMAIN 190 215 PORE-FORMING 2 (POTENTIAL).

FT TRANSSEM 216 250 POTENTIAL.

FT DOMAIN 251 325 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 77 77 N-LINKED (GLYCNIC...) (POTENTIAL).

SO SEQUENCE 499 AA; 55130 MW; B871A7A483DAD00 CRC64;

Query Match 74.0%; Score 37; DB 1; Length 499;

Best Local Similarity 75.0%; Pred. No. 7.1;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 ELISIPYE 11

DB 478 ELISIPYE 485

RESULT 7

POG_OMV

ID POG_OMV

AC P20234;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Genome polyprotein (containing: Nuclear inclusion protein A (N1-A) (N1A)

(EC 3.4.22.44) (49 Kda proteinase) (49 Kda-Pro); Nuclear inclusion

DE protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC 2.7.7.48);
 DE Coat protein (CP) (Ergament).
 OS Orithogalum mosaic virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage: Polyviridae;
 OC Polyvirus.
 RX NCBI_Taxid=12204;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Burger J.T., Brand R.J., Rybicki E.P.;
 RL Submitted (Nov-1990) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTOVIRAL ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
 CC further restricted by preferences for the amino acids in P6 - P1'
 CC that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
 CC Glu-Ser or Gly) for the enzyme from tobacco etch virus. The
 CC natural substrate is the viral polyprotein, but other proteins and
 CC oligopeptides containing the appropriate consensus sequence are
 CC also cleaved.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- PPM: THE VIRAL RNA OF POLYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POLYVIRUSES POLYPEPTIDE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D00615; BAA00490.1; .
 DR PIR: J00494; J00494.
 DR MEROPS: C04.005; .
 DR InterPro: IPR001730; Peptidase_C4.
 DR InterPro: IPR001592; Poly_coat.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00687; Poly_coat.1.
 DR Pfam: PF00863; Peptidase_C4.1.
 DR Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
 KW Coat protein; Polyprotein.
 FT NON_TER 1 1
 FT CHAIN <1 365 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 366 883 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 884 1135 COAT PROTEIN.
 SQ SEQUENCE 1136 AA; 128914 MW; 0A7AE0527743FD61 CRC64;

Query Match 74.0%; Score 37; DB 1; Length 1136;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 ELISPYEE 11
 Db 698 KLSIPEE 705

RESULT 8
 VEG_HPV31
 ID VEG_HPV31 STANDARD; PRT; 149 AA.
 AC P17386;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE E6 protein.
 GN E6.
 OS Human papillomavirus type 31.
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;

OC Papillomavirus.
 OC NCBI_Taxid=10585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89299478; PubMed=2545036;
 RA Goldsborough M.D., DiIvestre D., Temple G.F., Lorincz A.T.;
 RT "Nucleotide sequence of human papillomavirus type 31: a cervical
 RT neoplasia-associated virus.";
 RL Virology 171:306-311(1985).
 CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE ONCOGENIC POTENTIAL.
 CC OF THIS VIRUS (CERVICAL NEOPLASIA-ASSOCIATED VIRUS).
 CC -1- FUNCTION: EXHIBIT A STRONG, BUT NON SPECIFIC AFFINITY FOR DOUBLE
 CC STRANDED DNA (IN VITRO).
 CC -1- SUBCELLULAR LOCATION: Nuclear matrix associated.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J04353; AAA46950.1; .
 DR PIR: A32444; W6WJ31.
 DR InterPro: IPR001334; E6.
 DR Pfam: PF00518; E6.1.
 DR Early protein; DNA-binding; Nuclear protein; Zinc-finger;
 KW Oncogene.
 FT ZN_FING 30 66 POTENTIAL.
 FT ZN_FING 103 139 POTENTIAL.
 SQ SEQUENCE 149 AA; 17713 MW; 61D2A86C362767D9 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 149;
 Best Local Similarity 63.6%; Pred. No. 3;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
 Db 15 LSXALHPYDE 25

RESULT 9
 TH14_ASPOR
 ID TH14_ASPOR STANDARD; PRT; 327 AA.
 AC Q90U29;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thiazole biosynthetic enzyme, mitochondrial precursor.
 GN TH14 OR PT1A.
 OS Aspergillus oryzae.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Mitosporic Trichocomaceae; Aspergillus.
 RN NCBI_Taxid=5062;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HL1134;
 RX MEDLINE=20393355; PubMed=10945258;
 RA Kubodera T., Yamashita N., Nishimura A.;
 RT "Pyridoxamine resistance gene (ptra) of Aspergillus oryzae: cloning,
 RT characterization and application as a dominant selectable marker for
 RT transformation.";
 RL Biosci. Biotechnol. Biochem. 64:1416-1421(2000).
 CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF THE THIAMINE PRECURSOR
 CC THIAZOLE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (potential).
 CC -1- SIMILARITY: BELONGS TO THE TH14 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AF217503; AAF2544.1; -
 DR InterPro: IPR002922; Th14.
 DR Pfam: PF01946; Th14; 1.
 DR Trigrams: TIR002922; Th14; 1.
 KW Thiamine biosynthesis; Mitochondrion; Transit peptide; FAD; NAD.
 FT TRANSLIT 1
 FT CHAIN ? 327 MITOCHONDRION (POTENTIAL).
 FT NP_BIND 77 107 THIAZOLE BIOSYNTHETIC ENZYME.
 FT NP_BIND 77 107 FAD OR NAD (POTENTIAL).
 SQ SEQUENCE 327 AA; 35099 MW; 7C561EE06742B2AB CRC64;

Query Match 72.0%; Score 36; DB 1; Length 327;
 Best Local Similarity 75.0%; Pred. No. 7.1;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 ELISPYEE 11
 11:11111
 Db 138 ELISPYEE 145

RESULT 10
 ID NECL_MOUSE STANDARD; PRT; 753 AA.
 AC P21662; P22546;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neuroendocrine convertase 1 precursor (EC 3.4.21.93) (NEC 1) (PC1)
 DE (Prohormone convertase 1) (Proprotein convertase 1) (PC3) (Purin
 DE homology) (Propeptide processing protease).
 CN PCSK1 OR NECL OR NEC-1 OR ART-1.
 OS Mus musculus (Mouse), and
 OS Mus cookii (Cook's mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090, 10098;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE;
 RX MEDLINE=91319778; PubMed=1862107;
 RA Kornei J., Chun J., Harter D., Axel R.;
 RT "Isolation and functional expression of a mammalian prohormone
 RT processing enzyme, murine prohormone convertase 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:6834-6838(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE; STRAIN=IAP1;
 RX MEDLINE=92041727; PubMed=1657897;
 RA Nakayama K., Hosaka M., Hatsuzaawa K., Murakami K.;
 RT "Cloning and functional expression of a novel endoprotease involved
 RT in prohormone processing at dibasic sites.";
 RL J. Biochem. 109:803-806(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE;
 RX MEDLINE=91203919; PubMed=2017186;
 RA Seidah N.G., Marcinkiewicz M., Benjannet S., Gaspar L., Beaudien G.,
 RA Mellet M.-G., Lazure C., Mikiay M., Chretien M.;
 RT "Cloning and primary sequence of a mouse candidate prohormone
 RT convertase PC1 homologous to PC2, Furin, and Kex2: distinct
 RT chromosomal localization and messenger RNA distribution in brain and
 RT pituitary compared to PC2.";
 RL Mol. Endocrinol. 5:111-122(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-MOUSE; COOKII;
 RX MEDLINE=9110525; PubMed=1988934;
 RA Smeekens S.P., Avruch A.S., Iamendola J., Chan S.J., Steiner D.F.;
 RT "Identification of a CDNA encoding a second putative prohormone
 RT convertase related to PC2 in AtT20 cells and islets of Langerhans.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:340-344(1991).
 RN [5]
 RP SEQUENCE OF 214-478 FROM N.A.
 RC SPECIES-MOUSE; TISSUE=pituitary;
 RX MEDLINE=91000356; PubMed=2169760;
 RA Seidah N.G., Gaspar L., Mion P., Marcinkiewicz M., Mikiay M.,
 RA Chretien M.;
 RT "cDNA sequence of two distinct pituitary proteins homologous to Kex2
 RT and furin gene products: tissue-specific mRNAs encoding candidates
 RT for pro-hormone processing proteases.";
 RL DNA Cell Biol. 9:415-424(1990).
 RN [6]
 RP ERRATUM.
 RC SPECIES-MOUSE;
 RX MEDLINE=91090850; PubMed=2264933;
 RA Seidah N.G., Gaspar L., Mion P., Marcinkiewicz M., Mikiay M.,
 RA Chretien M.;
 RL DNA Cell Biol. 9:789-789(1990).
 CC -1- FUNCTION: INVOLVED IN THE PROCESSING OF HORMONE AND OTHER PROTEIN
 CC PRECURSORS AT SITES COMPRISED OF PAIRS OF BASIC AMINO ACID
 CC RESIDUES. SUBSTRATES INCLUDE POMC, RENIN, ENKEPHALIN, DYNORPHIN,
 CC SOMATOSTATIN AND INSULIN.
 CC -1- CATALYTIC ACTIVITY: Release of protein hormones, neuropeptides and
 CC renin from their precursors, generally by cleavage of -lys-Arg-1-
 CC bonds.
 CC -1- COFACTOR: CALCIUM DEPENDENT.
 CC -1- SUBCELLULAR LOCATION: LOCALIZED IN THE SECRETORY GRANULES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8. FURIN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M58507; AAA39896.1; -
 DR EMBL: X57088; CAA40368.1; ALT_SEQ.
 DR EMBL: M58589; AAA39894.1; -
 DR EMBL: M55668; AAA39375.1; ALT_SEQ.
 DR EMBL: M69196; AAA39732.1; -
 DR PIR: JK0171; KXMSC1.
 DR PIR: A37951; A37951.
 DR HSSP: Q45670; IDBI.
 DR MEROPS: S08.072; -
 DR MGD: MGI:97511; PCSK1.
 DR InterPro: IPR002884; P_domain.
 DR InterPro: IPR002094; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR Pfam: PF01483; P; PARTIAL.
 DR PRINTS: PR00723; SUBTILISIN.
 DR ProDom: PD000717; P_domain; 1.
 DR PROSITE: PS00136; SUBTILASE_AS1; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase; Serine protease; Glycoprotein; Zymogen; Calcium; Signal.
 FT SIGNAL 1 27
 FT PROPEP 28 110
 FT CHAIN 111 753
 FT DOMAIN 122 410
 FT DOMAIN 739 751
 FT ACT_SITE 167 167
 FT ACT_SITE 208 208
 FT ACT_SITE 382 382
 FT CARBOHYD 401 401
 FT CARBOHYD 645 645
 FT CONFLICT 23 23
 SQ SEQUENCE 753 AA; 84146 MW; 95B78441BE8BD9CD CRC64;

Query Match 72.0%; Score 36; DB 1; Length 753;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

FT CONFLICT 366 366 T -> TT (IN REF. 2).
 FT CONFLICT 514 514 E -> A (IN REF. 2).
 SQ SEQUENCE 752 AA; 84120 MW; F630AD830A0760ED CRC64;
 Query Match 70.0%; Score 35; DB 1; Length 752;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SKELTYPE 10
 1 :||:||||
 Db 689 SAKISPYE 697

RESULT 13
 DPOL_DEIRA STANDARD; PRT; 956 AA.
 AC P52027;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase I (EC 2.7.7.7) (POL I).
 GN POLA OR DRI107.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN 11
 RC SEQUENCE OF 36-956 FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=93273728; PubMed=8501062;
 RA Gultman P.D., Fuchs P., Ouyang L., Minton K.W.;
 RT "Identification, sequencing, and targeted mutagenesis of a DNA
 RT polymerase gene required for the extreme radioresistance of
 RT Deinococcus radiodurans";
 RL J. Bacteriol. 175:3581-3590(1993).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson J.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann J.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RI";
 RL Science 286:1571-1577(1999).
 CC -1- FUNCTION: IN ADDITION TO POLYMERASE ACTIVITY, THIS DNA POLYMERASE
 CC EXHIBITS 3' TO 5' AND 5' TO 3' EXONUCLEASE ACTIVITY (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + [DNA](N).
 CC -1- SUBUNIT: SINGLE-CHAIN MONOMER WITH MULTIPLE FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L14581; AAC36974.1; -
 CC EMBL: AEO02012; AAF1264.1; -
 CC HSSP: P19821; ZKTQ.
 DR TIGR: DRI107; -
 DR InterPro: IPR002421; 5_3_exonuclease.
 DR InterPro: IPR01098; DNA_Pol.
 DR InterPro: IPR002298; DNA_Pol.
 DR InterPro: IPR000513; Exo_N.I.

DR InterPro: IPR003583; HNH.1.
 DR InterPro: IPR003584; HNH.2.
 DR Pfam: PF00476; DNA_Pol_A; 1.
 DR Pfam: PF01367; 5_3_exonuclease; 1.
 DR Pfam: PF02739; 5_3_exonuclease; 1.
 DR PRINTS: PR00868; DNAPOL1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00278; HNH1; 1.
 DR SMART: SM00279; HNH2; 1.
 DR SMART: SM00482; POLA; 1.
 DR TIGRfams: TIGR00593; polA; 1.
 DR PROSITE: PS00447; DNA_POLYMERASE_A; 1.
 KW Hydrolase; Exonuclease; DNA-binding; complete proteome.
 KW Hydroxylase; Exonuclease; DNA-binding; complete proteome.
 FT CONFLICT 48 48 T -> A (IN REF. 1).
 FT CONFLICT 77 77 H -> D (IN REF. 1).
 FT CONFLICT 108 108 P -> L (IN REF. 1).
 FT CONFLICT 128 128 P -> R (IN REF. 1).
 FT CONFLICT 406 406 P -> R (IN REF. 1).
 FT CONFLICT 540 540 L -> V (IN REF. 1).
 SQ SEQUENCE 956 AA; 105659 MW; 6A8PF117075A884A CRC64;
 Query Match 70.0%; Score 35; DB 1; Length 956;
 Best Local Similarity 63.6%; Pred. No. 36;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SKELTYPE 11
 1 :||:||||
 Db 803 USNDIGIPYAB 813

RESULT 14
 TH14_YEAST STANDARD; PRT; 326 AA.
 AC P32518;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thiazole biosynthetic enzyme, mitochondrial precursor.
 GN TH14 OR MOL1 OR HSP35 OR YCR14W OR C6620.
 OS Saccharomyces cerevisiae (baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93070608; PubMed=1441749;
 RA Praekelt U.M., Meacock P.A.;
 RT "MOL1, a Saccharomyces cerevisiae gene that is highly expressed in
 RT early stationary phase during growth on molasses";
 RL Yeast 8:699-710(1992).
 RN 12
 RP SEQUENCE FROM N.A.
 RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Talla E.,
 RA Nawroth A., del Bino S., Goffeau A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN 13
 RP SEQUENCE OF 205-326 FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=96158062; PubMed=8585325;
 RA Skala J., Nawroth A., Goffeau A.;
 RT "The sequence of a 27 kb segment on the right arm of chromosome VII
 RT from Saccharomyces cerevisiae reveals MOL1, NAW2, RPL30B, NSR1, CYS4,
 RT PEM1/CH02, NSR1 genes and ten new open reading frames";
 RL Yeast 11:1421-1427(1995).
 RN 14
 RP THIAMINE REGULATION.
 RX MEDLINE=95028146; PubMed=7941734;
 RA Praekelt U.M., Byrne K.L., Meacock P.A.;
 RT "Regulation of TH14 (MOL1), a thiamine-biosynthetic gene of
 RT Saccharomyces cerevisiae";
 RL Yeast 10:481-490(1994).
 RN 15

RP FUNCTION IN DNA DAMAGE TOLERANCE.
 RX MEDLINE=98035046; PubMed=987751;
 RA Machado C.R., Prækel U.M., de Oliveira R.C., Barbosa A.C.,
 RA Byrne K.L., Meacock P.A., Menck C.F.;
 RT "Dual role for the yeast TH14 gene in thiamine biosynthesis and DNA
 RT damage tolerance";
 RT J. Mol. Biol. 273:114-121(1997).
 CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF THE THIAMINE PRECURSOR
 CC THIAZOLE. ALSO SEEMS TO HAVE A ROLE IN MITOCHONDRIAL DNA DAMAGE
 CC TOLERANCE.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (Potential).
 CC -1- INDUCTION: RERESSED BY THIAMINE.
 CC -1- MISCELLANEOUS: EXPRESSED AT HIGH LEVELS IN THE EARLY STATIONARY
 CC PHASE OF BATCH CULTURES GROWING ON MOLASSES, AN INDUSTRIAL MEDIUM.
 CC -1- SIMILARITY: BELONGS TO THE TH14 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X61669; CAA3843.1; -;
 DR EMBL: X85807; CAA59802.1; -;
 DR EMBL: 272929; CAA97157.1; -;
 DR EMBL: 272930; CAA97159.1; -;
 DR PIR: S17019; S17019.
 DR PIR: S25321; S25321.
 DR SGD: S0003376; TH14.
 DR InterPro: IPR002922; TH14.
 DR Pfam: PF01946; TH14; 1.*
 DR TIGRfams: TIGR00292; TH14; 1.
 DR Thiamine biosynthesis; Mitochondrion; Transit peptide; FAD: NAD.
 FT TRANSIT 1 ?
 FT CHAIN ? 326 THIAZOLE BIOSYNTHETIC ENZYME.
 FT NP_BIND 67 97 FAD OR NAD (POTENTIAL).
 SQ SEQUENCE 326 AA; 34991 MW; 843790F2CE08F02 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 326;
 Best Local Similarity 75.0%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELSPYEE 11
 Db 128 EMBLPIED 135
 Y103_METJA STANDARD; PRT; 433 AA.
 AC Q57567;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ0103.
 GN MJ0103.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: BELONGS TO THE UP0063 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67467; AA98083.1; -;
 DR TIGR: MJ0103; -;
 DR InterPro: IPR000385; MOA_NiFe_PqqE.
 DR Pfam: PF01444; MOA_NiFe_PqqE; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 433 AA; 49866 MW; F82576531DF12142 CRC64;

Query Match 68.0%; Score 34; DB 1; Length 433;
 Best Local Similarity 75.0%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELSPYEE 11
 Db 33 EMBLPIED 40

Search completed: January 15, 2003, 11:21:23
 Job time : 8.28572 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 10:29:39 ; Search time 25.5714 Seconds
(without alignments)
88.635 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 LSXELSTPYEE 11

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp Virus:*
16: sp Bacteriophage:*
17: sp Archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	80.0	1451	10 Q9X148	Q9X148 arbidopsin
2	39	78.0	634	16 Q44755	Q44755 arbidopsin
3	39	78.0	850	16 Q66691	Q66691 aquifer axo
4	38	76.0	474	16 Q80621	Q80621 agrobacteri
5	37	74.0	461	16 Q989F3	Q989F3 rhizobium I
6	37	74.0	465	12 Q9Q4W5	Q9Q4W5 ornithogaru
7	37	74.0	502	11 Q9JK62	Q9JK62 mus musculu
8	37	74.0	4717	3 Q9J248	Q9J248 schizosacch
9	36	72.0	291	16 Q8RB89	Q8RB89 thermomac
10	36	72.0	627	5 Q8SVF8	Q8SVF8 encephalit
11	36	72.0	1890	10 Q49528	Q49528 arbidopsin
12	35	70.0	208	16 Q9A3W3	Q9A3W3 caulobacter
13	35	70.0	274	5 Q9A849	Q9A849 drosophila
14	35	70.0	279	11 Q8VE10	Q8VE10 mus musculu
15	35	70.0	381	12 Q9YMM5	Q9YMM5 lymantiria d
16	35	70.0	697	4 Q8TC35	Q8TC35 homo sapien

17	35	70.0	738	5 Q9XW75	Q9XW75 caenorhabdi
18	35	70.0	1026	11 Q62845	Q62845 rattus norv
19	35	70.0	2894	17 Q58791	Q58791 methanococc
20	34	68.0	184	16 Q55139	Q55139 synchocyst
21	34	68.0	225	16 Q921Y9	Q921Y9 rickettsia
22	34	68.0	233	16 Q92XP6	Q92XP6 listeria in
23	34	68.0	233	16 Q8Y403	Q8Y403 listeria mo
24	34	68.0	235	17 Q9UX12	Q9UX12 sulfolobus
25	34	68.0	269	5 Q9N610	Q9N610 leishmania
26	34	68.0	359	16 Q92CJ1	Q92CJ1 listeria in
27	34	68.0	359	16 Q8Y705	Q8Y705 listeria mo
28	34	68.0	417	17 Q9V7T1	Q9V7T1 pyrococcus
29	34	68.0	428	17 Q8T1S8	Q8T1S8 methanocarc
30	34	68.0	577	10 Q9V7X2	Q9V7X2 oryza sativ
31	34	68.0	597	3 Q59717	Q59717 schizosacch
32	34	68.0	697	17 Q8TJ04	Q8TJ04 melanosarc
33	34	68.0	1028	4 Q9T052	Q9T052 homo sapien
34	34	68.0	1028	11 Q9JMH8	Q9JMH8 mus musculu
35	34	68.0	1028	11 Q9JMH8	Q9JMH8 rattus norv
36	34	68.0	1085	5 Q24363	Q24363 drosophila
37	34	68.0	1234	5 Q24690	Q24690 drosophila
38	34	68.0	6815	5 Q91704	Q91704 drosophila
39	34	68.0	16215	5 Q9NE53	Q9NE53 drosophila
40	33	66.0	118	16 P72904	P72904 synchocyst
41	33	66.0	132	10 Q9T0G2	Q9T0G2 arbidopsin
42	33	66.0	153	16 Q926Q7	Q926Q7 listeria in
43	33	66.0	153	16 Q8Y314	Q8Y314 listeria mo
44	33	66.0	199	4 Q9NMM6	Q9NMM6 homo sapien
45	33	66.0	228	4 Q96124	Q96124 homo sapien

ALIGNMENTS

RESULT 1	ID	Q9X148	PRELIMINARY	PRT	1451 AA.
AC	Q9X148				
DT	01-NOV-1999 (TREMBLrel. 12, Created)				
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	P9J1.15 protein.				
GN	P9J1.15.				
OS	Arbidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. COLUMBIA;				
RA	Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lenz G., Liu S.,				
RA	Lee J., Liu A., Li J., Kremenetskaia I., Juros J., Gonzalez A.,				
KA	Altef H., Araujo K., Brooks S., Buchler E., Chao Q., Conn L.,				
KA	Conway A.B., Dunn P., Hansen N., Hotz R., Khan S., Kim C., Palm C.,				
RA	Rowley D., Shin P., Walker M., Davis R.W., Ecker J.R.,				
RT	"Arbidopsis thaliana chromosome 1 BAC F9L1 sequence."				
RL	Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.				
EMBL	AC007591; AA039650.1; -				
DR	InterPro: IPR003593; AAA_AITase.				
DR	InterPro: IPR003439; ABC_transport.				
DR	Plan: PF00005; ABC_tran; 2.				
DR	ProDom: PD000006; ABC_transport; 1.				
DR	SMART: SM00382; AAA; 1.				
KW	ATP-binding.				
SO	SEQUENCE 1451 AA; 163602 MW; 0D35414D9C370A85 CRC64;				
Query Match	80.0%;				
Best local similarity	63.6%;				
Matches	7; Conservative				
	3; Mismatches				
	1; Indels				
	Gaps				
	0;				
QY	1 LSXELSTPYEE 11				

Db 491 LSXELSVYDK 501

RESULT 2

044755

PRELIMINARY;

PRT; 634 AA.

AC 044755;

DT 01-NOV-1996 (TRENBLREL. 01, Created)

DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)

DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)

DE ORF37.

GN B80257.

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 35210 / B31;

RL Dunn J.J., Butler-Loffredo L., Kleczawa J., Medalle J., Luft B.J., Submitted (DRC-1995) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 35210 / B31;

RE MEDLINE=98065943; PubMed=9403685;

RA Fraser C.M., Castjens S., Huang W.M., Sutton G.G., Clayton R.A.,

RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwin M.,

RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,

RA Peterson J., Kertavage A.R., Quackenbush J., Salzberg S., Hanson M.,

RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

RA Uterback T., Wathley L., McDonald L., Atlach P., Bowman C.,

RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,

RA Smith H.O., Venter J.C.,

RT "Genomic sequence of a Lyme disease spirochete, Borrelia

burgdorferi."

RL Nature 380:580-586(1997).

DR EMBL; U43739; AA85591.1; -

DR EMBL; AE001137; AAC6681.1; -

DR TIGR; B80267; -

KW Complete proteome.

SO SEQUENCE

634 AA; 70752 MW; E8AFB112B121F3B8 CRC64;

QY

2 SXELSVYDK 11

Db 43 SLELSVYDK 52

Query Match

Best Local Similarity

Matches

7; Conservative

2; Mismatches

1; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 634;

78.08;

70.08;

Pred. No. 23;

1; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 634;

78.08;

70.08;

Pred. No. 23;

1; Indels

0; Gaps

0;

DR EMBL; AE000686; AAC06656.1; -

DR InterPro; IPR001633; EAL.

DR InterPro; IPR001633; GDEF.

DR InterPro; IPR001633; Mitochondrion.

DR Pfam; PF00563; EAL; 1.

DR Pfam; PF00990; GDEF; 1.

DR SMART; SM00267; DUF2.1.

DR SMART; SM00552; DUF2.1.

DR TIGR; TIGR00254; GDEF; 1.

DR PROSITE; PS00215; MITOCHONDRIAL_CARRIER.

DR Hypothetical protein; Complete proteome.

KW SEQUENCE

850 AA; 97669 MW; C02DCEP91DCEB0 CRC64;

Query Match

Best Local Similarity

Matches

6; Conservative

3; Mismatches

2; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 850;

78.08;

54.58;

Pred. No. 31;

2; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 850;

78.08;

54.58;

Pred. No. 31;

2; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 850;

78.08;

54.58;

Pred. No. 31;

2; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 850;

78.08;

54.58;

Pred. No. 31;

2; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 850;

78.08;

54.58;

Pred. No. 31;

2; Indels

0; Gaps

0;

Score 39;

DB 16;

Length 850;

78.08;

54.58;

Pred. No. 31;

2; Indels

0; Gaps

0;

Db 200 LAMDDIPIYE 210

RESULT 5

Q989F3

1: :|:|:|:|:|:|
ID Q989F3 PRELIMINARY; PRT: 461 AA.
AC Q989F3;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Alpha-galactosidase.
GN MR6450.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Pyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAP303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsumo A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003009; BAB52744.1; -;
DR InterPro: IPR001089; GH_4.
DR Pfam: PF02056; Glyco_hydro.4; 1.
DR PRINTS: PR00732; GLHYDRLS4.
DR PRODOM: PD006892; GH_4; 1.
DR PROSITE: PS01324; GLYCOSTL_HYDROL_F4; UNKNOWN_1.
RW Complete proteome.
SQ SEQUENCE 461 AA; 51818 MW; 6C29F18157E4F6EC CRC64;

Query Match 74.0%; Score 37; DB 16; Length 461;
Best local Similarity 54.5%; Pred. No. 40;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 11
1: :|:|:|:|:|:|
Db 188 LAMDDIPIYE 198

RESULT 6

Q9Q4M5

PRELIMINARY; PRT: 465 AA.
AC Q9Q4M5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyprotein (Fragment).
OS Ornithogalum moschicola virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Polyvirus.
OX NCBI_TaxID=12204;
RN [1]
RP SEQUENCE FROM N.A.
RC Mackenzie A., Gibbs A.;
RT "Polyviruses in Australia.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF203528; AAF22752.1; -;
DR InterPro: IPR001592; Poly_coat.
DR Pfam: PF00767; Poly_coat; 1.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
FT MON_TER 1
SQ SEQUENCE 465 AA; 53264 MW; 3D4B484C359FF636 CRC64;

Query Match 74.0%; Score 37; DB 12; Length 465;

Best local Similarity 87.5%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 ELISPIEE 11
1: :|:|:|:|:|:|
Db 27 KLSIPIEE 34

RESULT 7

Q9JK62

PRELIMINARY; PRT: 502 AA.
AC Q9JK62;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Potassium channel TASK2 (TASK2 potassium channel).
GN KCKN5.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ; TISSUE=KIDNEY;
RA Roux J., Barhanin J.;
RT "Mouse two P domain potassium channel TASK2.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Cid L.P., Memey M.J., Sepulveda P.V.;
RT "Functional properties of mouse TASK-2 potassium channel.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF259395; AAF68668.1; -;
DR EMBL: AF319542; AAC35065.1; -;
DR MGD: MGI:136175; Kcnk5.
DR InterPro: IPR003280; Kchannel_2pore.
DR InterPro: IPR001622; Kchannel_pore.
DR PRINTS: PRO1333; ZP0REKCHANEL.
RW Ionic channel.
SQ SEQUENCE 502 AA; 55976 MW; EAC7E7CC71B4C075 CRC64;

Query Match 74.0%; Score 37; DB 11; Length 502;
Best local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 ELISPIEE 11
1: :|:|:|:|:|:|
Db 481 ELISPIEE 488

RESULT 8

Q94248

PRELIMINARY; PRT: 4717 AA.
AC Q94248;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical coiled-coil protein.
GN SPOC737.08.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Murphy L., Harris D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;

Query Match 74.0%; Score 37; DB 12; Length 465;

QY 2 SXELSIPEE 11
11111111
DB 1137 SLEESVPEE 1146

RESULT 12

Q9A3W3 PRELIMINARY; PRT; 208 AA.

AC Q9A3W3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glutathione S-transferase family protein.
GN CC3088.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
CC Caulobacter.
OX NCBI_TaxID=155892;

RM [1]
RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed=11259647;
RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.F., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Padke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Hafi D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ueberlack T., Tran K., Wolf A., Yamalvean J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

DR EMBL: AE005972; AAK25050+1; -
DR TIGR: CC3088; -
DR InterPro: IPR004046; GST_Cterm.
DR InterPro: IPR004045; GST_Nterm.
DR Pfam: PF02043; GST_C; 1.
DR Pfam: PF02798; GST_N; 1.
KV Transferase; Complete proteome.
SQ SEQUENCE 208 AA; 22885 MW; A48AC2988C284DD1 CRC64;

Query Match 70.0%; Score 35; DB 16; Length 208;
Best Local Similarity 70.0%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 10
11111111
DB 18 LLELSLPEE 27

RESULT 13

Q94849 PRELIMINARY; PRT; 274 AA.

AC Q94849;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Yolk protein 1 (Fragment).
GN YP1.
OS Drosophila buzzatii (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7264;

RM [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96120851; PubMed=8548285;
RA Kambsvelli M.P., Ho K.F., Craddock E.M., Piano F., Parisi M.,
RA Cohen J.;
RT "Pattern of ecological shifts in the diversification of Hawaiian
RT Drosophila inferred from a molecular phylogeny."
RL Curr. Biol. 5:1129-1139(1995).

RM [2]
RP Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RA Ho K.F.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U52126; AAC47249.1; -
DR FlyBase: FBgn0016167; DbuzYp1.
DR InterPro: IPR000734; Lipase.
DR Pfam: PF00151; Lipase; 1.
FT NON_TER 1 1
FT NON_TER 274 274
SQ SEQUENCE 274 AA; 30637 MW; 772CFE9AE71BB48 CRC64;

Query Match 70.0%; Score 35; DB 5; Length 274;
Best Local Similarity 60.0%; Pred. No. 59;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSXELSIPEE 10
11111111
DB 220 LMSLESLPEE 229

RESULT 14

Q8VE10 PRELIMINARY; PRT; 279 AA.

AC Q8VE10;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical 32.1 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RM [1]
RP SEQUENCE FROM N.A.
RA Stensberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
KV Hypothetical protein.
SQ SEQUENCE 279 AA; 32067 MW; 4AA9EB45DAB21A01 CRC64;

Query Match 70.0%; Score 35; DB 11; Length 279;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 ELISPEE 11
11111111
DB 18 ELISPEE 25

RESULT 15

Q9YMM5 PRELIMINARY; PRT; 381 AA.

AC Q9YMM5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ldori-102 peptide.
OS Lymnaea dispar multicapsid nuclear polyhedrosis virus (LdMPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10449;

RM [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99124785; PubMed=9887315;
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J.M., Rohmann G.F.;
RT "Sequence and analysis of the genome of a baculovirus pathogenic for
RT Lymnaea dispar."
RL Virology 253:17-34(1999).

RM [2]
RP SEQUENCE FROM N.A.
RP Kuzio J.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF081810; AAC70288.1; -
SQ SEQUENCE 381 AA; 43418 MW; EA63951C43C3ACA6 CRC64;

Query Match 70.0%; Score 35; DB 12; Length 381;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ELSIPYEE 11
||:||||
Db 41 ELDLPYEE 48

Search completed: January 15, 2003, 12:35:48
Job time : 26.5714 secs

GenoCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 16, 2003, 08:03:43 ; Search time 1525.14 Seconds

(without alignments)
209,902 Million cell updates/sec

Title: US-09-823-649A-4
Perfect score: 50
Sequence: 1 ISXELSIPTPE 11

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters: --
-MODE=frame+p2n,model-DEV=xlh
-DB=genmb1,USPTO.spool/US09823649/runat.14012003.151002.29139/app.query.fasta.1.1393
-DB=genmb1,OPMT=fastap,SUFFIX=rge-MINMATCH=0.1-LOOPEL=0-LOOPEXT=0
-UNITS=bits-START=1-END=1-MATRIX=blousum62-TRANS=human40.cnt-List=45
-DOCALLIGN=200-THR.SCORE=pct-THR.MAX=100-THR.MIN=0-ALIGN=15-MODE=LOCAL
-OUTFMT=ptc-NORM=ext-HEAPSIZ=500-MINLEN=0-MAXLEN=2000000000
-USER=US09823649.@CGN.1.1_3378.@runat.14012003.151002.29139-MCPU=6-ICPU=3
-NO.XLPHY-NO.MAP-LARGSEQUERY-NEG.SCORES=0-WAIT-LONGLOG-DEV.TIMEOUT=120
-WARN.TIMEOUT=30-THREADS=1-XGAPOP=10-XGAPEXT=0.5-FGAPOP=6-FGAPEXT=7
-YGAPOP=10-YGAPEXT=0.5-DELOP=6-DELEXT=7

Database : Genmb1:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vtl:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.hugo.other:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	48	96.0	2493 6 AR023936	AR023936 Sequence
2	48	96.0	2493 6 I15437	I15437 Sequence 5
3	48	96.0	2496 6 AR023961	AR023961 Sequence
4	48	96.0	2496 6 AR054698	AR054698 Sequence
5	48	96.0	2496 6 AR061694	AR061694 Sequence
6	48	96.0	2496 6 AR061935	AR061935 Sequence
7	48	96.0	2496 6 AR063636	AR063636 Sequence
8	48	96.0	2496 6 AR086117	AR086117 Sequence
9	48	96.0	2496 6 AR089547	AR089547 Sequence
10	48	96.0	2496 6 AR093632	AR093632 Sequence
11	48	96.0	2496 6 AR193571	AR193571 Sequence
12	48	96.0	2496 6 AR206169	AR206169 Sequence
13	48	96.0	2496 6 AR317151	AR317151 Sequence
14	48	96.0	2496 6 I24091	I24091 Sequence 2
15	48	96.0	2496 6 I24091	I24091 Sequence 2
16	48	96.0	2496 6 I24091	I24091 Sequence 2
17	48	96.0	3403 1 TPOLDNA	TPOLDNA Sequence 2
18	45	90.0	1079 8 AF315118	AF315118 Sequence
19	45	90.0	1600 6 AR023967	AR023967 Sequence
20	45	90.0	1600 6 AR054704	AR054704 Sequence
21	45	90.0	1600 6 AR061700	AR061700 Sequence
22	45	90.0	1600 6 AR061941	AR061941 Sequence
23	45	90.0	1600 6 AR063642	AR063642 Sequence
24	45	90.0	1600 6 AR086123	AR086123 Sequence
25	45	90.0	1600 6 AR089553	AR089553 Sequence
26	45	90.0	1600 6 AR093638	AR093638 Sequence
27	45	90.0	1600 6 AR193577	AR193577 Sequence
28	45	90.0	1600 6 AR206175	AR206175 Sequence
29	45	90.0	1600 6 I24097	I24097 Sequence 12
30	45	90.0	1600 6 I38585	I38585 Sequence 12
31	45	90.0	1600 6 I76140	I76140 Sequence 12
32	45	90.0	1682 6 AR112315	AR112315 Sequence
33	45	90.0	1696 6 AR112315	AR112315 Sequence
34	45	90.0	1812 6 AX093143	AX093143 Sequence
35	45	90.0	1812 6 AX093143	AX093143 Sequence
36	45	90.0	2433 6 AX317584	AX317584 Sequence
37	45	90.0	2445 6 AX317568	AX317568 Sequence
38	45	90.0	2445 6 AX317572	AX317572 Sequence
39	45	90.0	2445 6 AX317576	AX317576 Sequence
40	45	90.0	2445 6 AX317578	AX317578 Sequence
41	45	90.0	2480 6 AX404968	AX404968 Sequence
42	45	90.0	2483 6 AX404966	AX404966 Sequence
43	45	90.0	2484 6 AX404972	AX404972 Sequence
44	45	90.0	2493 6 AX317586	AX317586 Sequence
45	45	90.0	2496 6 AX317344	AX317344 Sequence

ALIGNMENTS

RESULT 1

AR023936
LOCUS AR023936 2493 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 5 from patent US 5795762.
ACCESSION AR023936
VERSION AR023936.1 GI:3977230
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE
1 (bases 1 to 2493)
AUTHORS Abramson, R.D. and Gelfand, D.H.
TITLE 5' to 3' exonuclease mutations of thermostable DNA polymerases
JOURNAL Patent: US 5795762-A 5 18-AUG-1998;
FEATURES
Location/Qualifiers
source
1..2493
BASE COUNT 411 a 859 c 843 g 380 t
ORIGIN

Alignment Scores:
Pred. No.: 1.27 Length: 2493
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR023936 (1-2493)

OY 1 LeuSer**GluLeuSerIleProTyrGluGln 11
LOCUS AR023961 2496 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 2 from patent US 5795763.
ACCESSION AR023961
VERSION AR023961.1 GI:3977255

Alignment Scores:
Pred. No.: 1.27 Length: 2493
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x I15437 (1-2493)

OY 1 LeuSer**GluLeuSerIleProTyrGluGln 11
LOCUS I15437 2493 bp DNA linear PAT 02-APR-1996
DEFINITION Sequence 5 from patent US 546551.
ACCESSION I15437
VERSION I15437.1 GI:1250345
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE
1 (bases 1 to 2493)
AUTHORS Abramson, R.D. and Gelfand, D.H.
TITLE 5' to 3' exonuclease mutations of thermostable DNA polymerases
JOURNAL Patent: US 546551-A 5 14-NOV-1995;
FEATURES
Location/Qualifiers
source
1..2493
BASE COUNT 411 a 859 c 843 g 380 t
ORIGIN

KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE
1 (bases 1 to 2496)
AUTHORS Dahlberg, J.F., Lyamichev, V.I. and Brown, M. Ann. D.
TITLE Synthesis-deficient thermostable DNA polymerase
JOURNAL Patent: US 5795763-A 2 18-AUG-1998;
FEATURES
Location/Qualifiers
source
1..2496
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN

Alignment Scores:
Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR023961 (1-2496)

OY 1 LeuSer**GluLeuSerIleProTyrGluGln 11
LOCUS AR054698 2496 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5837450.
ACCESSION AR054698
VERSION AR054698.1 GI:5980275
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE
1 (bases 1 to 2496)
AUTHORS Dahlberg, J.F., Lyamichev, V.I. and Brown, M. Ann. D.
TITLE Detection of target nucleic acid molecules using thermostable 5' nuclease
JOURNAL Patent: US 5837450-A 2 17-NOV-1998;
FEATURES
Location/Qualifiers
source
1..2496
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN

Alignment Scores:
Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR054698 (1-2496)

OY 1 LeuSer**GluLeuSerIleProTyrGluGln 11
LOCUS AR061694 2496 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5843654.
ACCESSION AR061694
VERSION AR061694.1 GI:5989385
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE
1 (bases 1 to 2496)
AUTHORS Dahlberg, J.F., Lyamichev, V.I. and Brown, M. Ann. D.
TITLE Synthesis-deficient thermostable DNA polymerase
JOURNAL Patent: US 5843654-A 2 17-NOV-1998;
FEATURES
Location/Qualifiers
source
1..2496
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN

REFERENCE 1 (bases 1 to 2496)
AUTHORS Heister/L.M., FOS/L. and Brow.M,Ann.D.
TITLE Rapid detection of mutations in the p53 gene
JOURNAL Patent: US 5843654-A 2 01-DEC-1998;
FEATURES Location/Qualifiers
source 1..2496
/organism="unknown"
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN
Alignment Scores:
Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-4 (1-11) x AR061694 (1-2496)
QY 1 Leuser***GluLeuser1leProTyGluGlu 11
Db 2029 CTCGCGGAGGACTTTCATCCCTACGAGAG 2061
RESULT 6 AR061935 2496 bp DNA linear PAT 29-SEP-1999
LOCUS AR061935
DEFINITION Sequence 2 from patent US 5843669.
ACCESSION AR061935
VERSION AR061935.1 GI:5989626
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Kaiser,M.W., Lyamichev,V.I. and Lyamichev,N.
TITLE Cleavage of nucleic acid using thermostable methanococcus
JOURNAL Patent: US 5843669-A 2 01-DEC-1998;
FEATURES Location/Qualifiers
source 1..2496
/organism="unknown"
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN
Alignment Scores:
Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-4 (1-11) x AR061935 (1-2496)
QY 1 Leuser***GluLeuser1leProTyGluGlu 11
Db 2029 CTCGCGGAGGACTTTCATCCCTACGAGAG 2061
RESULT 7 AR063636 2496 bp DNA linear PAT 29-SEP-1999
LOCUS AR063636
DEFINITION Sequence 2 from patent US 5846717.
ACCESSION AR063636
VERSION AR063636.1 GI:5992944
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Brow,M,Ann.D., Hall,J,Steven,Goetelueschen., Lyamichev,V.,
Olive,D,Michael. and Prudent,J,Robert.
TITLE Detection of nucleic acid sequences by invader-directed cleavage

JOURNAL Patent: US 5846717-A 2 08-DEC-1998;
FEATURES Location/Qualifiers
source 1..2496
/organism="unknown"
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN
Alignment Scores:
Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-4 (1-11) x AR063636 (1-2496)
QY 1 Leuser***GluLeuser1leProTyGluGlu 11
Db 2029 CTCGCGGAGGACTTTCATCCCTACGAGAG 2061
RESULT 8 AR066117 2496 bp DNA linear PAT 07-SEP-2006
LOCUS AR066117
DEFINITION Sequence 2 from patent US 5985557.
ACCESSION AR066117
VERSION AR066117.1 GI:10012883
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Prudent,J,R., Hall,J,G., Lyamichev,V.I., Brow,M,Ann.D. and
Dahlberg,J,E.
TITLE Invasive cleavage of nucleic acids
JOURNAL Patent: US 5985557-A 2 16-NOV-1999;
FEATURES Location/Qualifiers
source 1..2496
/organism="unknown"
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN
Alignment Scores:
Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0
US-09-823-649A-4 (1-11) x AR066117 (1-2496)
QY 1 Leuser***GluLeuser1leProTyGluGlu 11
Db 2029 CTCGCGGAGGACTTTCATCCCTACGAGAG 2061
RESULT 9 AR089547 2496 bp DNA linear PAT 07-SEP-2006
LOCUS AR089547
DEFINITION Sequence 2 from patent US 5994069.
ACCESSION AR089547
VERSION AR089547.1 GI:10016304
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Hall,J,G., Lyamichev,V.I., Mast,A,L. and Brow,M,Ann.D.
TITLE Detection of nucleic acids by multiple sequential invasive
cleavages
JOURNAL Patent: US 5994069-A 2 30-NOV-1999;
FEATURES Location/Qualifiers
source 1..2496

BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN /organism="unknown"

Alignment Scores:

Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR089547 (1-2496)

QY 1 LeuSer***GluLeuSerIleProTyrGluGlu 11

Db 2029 CTCCTCCGGGAGCTTCCATCCCTACGAGAG 2061

RESULT 10

LOCUS AR093632 2496 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 2 from patent US 6001567.

ACCESSION AR093632

VERSION / 091632.1 GI:10020381

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1.27 Length: 2496

Score: 48.00 Matches: 10

Percent Similarity: 90.91% Conservative: 0

Best Local Similarity: 90.91% Mismatches: 1

Query Match: 96.00% Indels: 0

DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR093632 (1-2496)

QY 1 LeuSer***GluLeuSerIleProTyrGluGlu 11

Db 2029 CTCCTCCGGGAGCTTCCATCCCTACGAGAG 2061

RESULT 11

LOCUS AR193571 2496 bp DNA linear PAT 20-APR-2002

DEFINITION Sequence 2 from patent US 6348314.

ACCESSION AR193571

VERSION AR193571.1 GI:20240163

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1.27 Length: 2496
Score: 48.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 96.00% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR193571 (1-2496)

QY 1 LeuSer***GluLeuSerIleProTyrGluGlu 11

Db 2029 CTCCTCCGGGAGCTTCCATCCCTACGAGAG 2061

RESULT 12

LOCUS AR206169 2496 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 2 from patent US 6372424.

ACCESSION AR206169

VERSION AR206169.1 GI:21504691

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1.27 Length: 2496

Score: 48.00 Matches: 10

Percent Similarity: 90.91% Conservative: 0

Best Local Similarity: 90.91% Mismatches: 1

Query Match: 96.00% Indels: 0

DB: 6 Gaps: 0

US-09-823-649A-4 (1-11) x AR206169 (1-2496)

QY 1 LeuSer***GluLeuSerIleProTyrGluGlu 11

Db 2029 CTCCTCCGGGAGCTTCCATCCCTACGAGAG 2061

RESULT 13

LOCUS AX317151 2496 bp DNA linear PAT 14-DEC-2001

DEFINITION Sequence 154 from patent W0190337.

ACCESSION AX317151

VERSION AX317151.1 GI:17900145

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

ORIGIN
Alignment Score: 1.27
Pred. No.: 48.00
Score: 48.00
Percent Similarity: 90.91%
Best Local Similarity: 90.91%
Query Match: 96.00%
DB: 6
Gaps: 0

US-09-823-649A-4 (1-11) x AX317151 (1-2496)

QY 1 LeuSer**GluLeuSerIleProTyrgluGlu 11
||||| ||||||| ||||||| ||||||| |||||||
Db 2029 CTCCTCGGGAGCTTCCATCCCTACGAGAG 2061

RESULT 14
LOCUS 124091 2496 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 2 from patent US 5541311.
ACCESSION 124091
VERSION 124091.1 GI:1603961
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Dahlberg,J.E., Lyamichev,V.I. and Brow,M.Ann.D.
TITLE Nucleic acid encoding synthesis-deficient thermostable DNA
polymerase
JOURNAL Patent: US 5541311-A 2 30-JUL-1996;
FEATURES Location/Qualifiers
source 1..2496
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN

Alignment Scores:
Pred. No.: 1.27
Score: 48.00
Percent Similarity: 90.91%
Best Local Similarity: 90.91%
Query Match: 96.00%
DB: 6
Gaps: 0

US-09-823-649A-4 (1-11) x 124091 (1-2496)

QY 1 LeuSer**GluLeuSerIleProTyrgluGlu 11
||||| ||||||| ||||||| ||||||| |||||||
Db 2029 CTCCTCGGGAGCTTCCATCCCTACGAGAG 2061

RESULT 15
LOCUS 138579 2496 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 2 from patent US 5614402.
ACCESSION 138579
VERSION 138579.1 GI:2084633
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2496)
AUTHORS Dahlberg,J.E., Lyamichev,V.I. and Brow,M.Ann.D.
TITLE 5' nucleases derived from thermostable DNA polymerase
JOURNAL Patent: US 5614402-A 2 25-MAY-1997;
FEATURES Location/Qualifiers
source 1..2496
BASE COUNT 402 a 830 c 854 g 410 t
ORIGIN

Alignment Scores:
Pred. No.: 1.27
Score: 48.00
Percent Similarity: 90.91%
Best Local Similarity: 90.91%
Query Match: 96.00%
DB: 6
Gaps: 0

Score: 48.00
Percent Similarity: 90.91%
Best Local Similarity: 90.91%
Query Match: 96.00%
DB: 6
Gaps: 0

US-09-823-649A-4 (1-11) x 138579 (1-2496)

QY 1 LeuSer**GluLeuSerIleProTyrgluGlu 11
||||| ||||||| ||||||| ||||||| |||||||
Db 2029 CTCCTCGGGAGCTTCCATCCCTACGAGAG 2061

Search completed: January 16, 2003, 11:02:27
Job time: 1530.14 secs

Matches: 10
Conservative: 0
Mismatch: 1
Indels: 0
Gaps: 0

DR WPI; 2002-076891/11.
XX Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermocactive DNA polymerase
PS Claim 4; Page 4; 23pp; English.
XX The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermocactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction. This motif is derived from DNA polymerases
CC from *Thermus* species flavus, sps17 and filliformis.
SO Sequence 11 AA:
Query Match 96.0%; Score 48; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSXELSIPEE 11
DB 1 LSXELSIPEE 11
RESULT 2
AAM48260
ID AAM48260 standard; Peptide; 11 AA.
XX
AC AAM48260;
XX
DT 25-MAR-2002 (first entry)
XX
DE Native DNA polymerase motif #9.
XX
KM DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KM transverse transcription reaction; fluorescein; cyanine; thermocactive;
KM dye; amplification.
XX
OS *Thermus* flavus.
XX
PN EPI152062-A2.
XX
PD 07-NOV-2001.
XX
PF 12-APR-2001; 2001EP-0109341.
XX
PR 18-APR-2000; 2000US-198336P.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
XX
DR WPI; 2002-076891/11.
XX Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermocactive DNA polymerase
PS Disclosure; Page 7; 23pp; English.
XX

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermocactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.
SO Sequence 11 AA:
Query Match 96.0%; Score 48; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0051;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSXELSIPEE 11
DB 1 LSXELSIPEE 11
RESULT 3
AAM48263
ID AAM48263 standard; Peptide; 11 AA.
XX
AC AAM48263;
XX
DT 25-MAR-2002 (first entry)
XX
DE Native DNA polymerase motif #12.
XX
KM DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KM transverse transcription reaction; fluorescein; cyanine; thermocactive;
KM dye; amplification.
XX
OS *Thermus* species sps17.
XX
PN EPI152062-A2.
XX
PD 07-NOV-2001.
XX
PF 12-APR-2001; 2001EP-0109341.
XX
PR 18-APR-2000; 2000US-198336P.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
XX
DR WPI; 2002-076891/11.
XX Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PT using a mutant thermocactive DNA polymerase
PS Disclosure; Page 7; 23pp; English.
XX The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of motifs derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermocactive DNA
CC polymerase. These motifs represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method

CC provides improved reverse transcription efficiency at lower enzyme
 CC concentrations relative to previous high temperature reverse
 CC transcription methods, and provide a wider range of usable salt
 CC concentrations. The use of the mutant DNA polymerase provides faster
 CC reverse transcription extension rates, and consequently less time is
 CC needed for the reaction.

XX Sequence 11 AA;

Query Match 96.0%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSPYEE 11
 || |||||
 DB 1 LSQELSPYEE 11

RESULT 4

AA048265
 ID AA048265 standard; Peptide; 1) AA.

XX AA048265;

DT 25-MAR-2002 (first entry)

DE Native DNA polymerase motif #14.

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;
 KW transverse transcription reaction; fluorescein; cyanine; thermostable;
 KW dye; amplification.

OS Thermus filiformis. -- *

PN EPI152062-A2.

XX 07-NOV-2001.

PF 12-APR-2001; 2001EP-0109341.

PR 18-APR-2000; 2000US-198336P.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers JW;

PI Schoenbrunner NJ, Wang AM;

DR WPI: 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse
 PT transcriptase polymerase chain reaction amplification of a mixture
 PT using a mutant thermostable DNA polymerase

PS Disclosure; Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AA048270 represent the
 CC native forms of motifs derived from DNA polymerases used in the method
 CC of the invention. The method for reverse transcribing an RNA, comprises
 CC treating a transverse transcription reaction mixture which comprises the
 CC RNA, a primer, a divalent cation, and a mutant thermostable DNA
 CC polymerase. These motifs represent a conserved region which affects the
 CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
 CC fluorescein and cyanine family dyes. The method of the invention is
 CC useful in reverse transcription/amplification reactions. The method
 CC provides improved reverse transcription efficiency at lower enzyme
 CC concentrations relative to previous high temperature reverse
 CC transcription methods, and provide a wider range of usable salt
 CC concentrations. The use of the mutant DNA polymerase provides faster
 CC reverse transcription extension rates, and consequently less time is
 CC needed for the reaction.

XX Sequence 11 AA;

Query Match 96.0%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0051;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSXELSPYEE 11
 || |||||
 DB 1 LSQELSPYEE 11

RESULT 5

AA023157
 ID AA023157 standard; Protein; 545 AA.

XX AA023157;

DT 22-OCT-1992 (first entry)

DE Mutant thermostable DNA polymerase enzyme MET-ALA 288 Tps17.

KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PCR.

OS Thermus species SPS17.

FN Key Location/Qualifiers

FT MISC-difference 1..2 /note- "residues 2-287 deleted from the native
 FT sequence"

PN MO9206200-A.

PD 16-APR-1992.

PF 30-SHP-1991; 91WO-US07035.

PR 28-SHP-1990; 90US-0590213.

PR 28-SEP-1990; 90US-0590466.

PR 28-SEP-1990; 90US-0590490.

PA (CENTU) CENTU CORP.

PI Abramson RD, Gelfand DH;

DR WPI: 1992-150885/18.

DR N-PSDB: AA024010.

PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays

PS Claim 11; Page 59; 185pp; English.

CC The sequence is that of a mutant of Thermus species SPS17 polymerase
 CC mutant MET-ALA 288 Tps17, having a different amt. of 5'-3' exonuclease
 CC activity than the native enzyme. Thermostable DNA polymerases
 CC are useful in many recombinant DNA techniques, esp. nucleic acid
 CC amplification by PCR, self-sustained sequence replication (SSR)
 CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
 CC may facilitate higher sensitivity allelic discrimination in a combined
 CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
 CC 5'-3' exonuclease activity may be desirable in enzymes used in
 CC homogeneous assays for the amplification and detection of a target
 CC nucleic acid sequence. Mutation of the DNA encoding particular
 CC regions of the enzymes can be used to prepare a range of recombinant
 CC proteins having 5'-3' exonuclease activity to a complete lack of
 CC activity.
 CC See also AAR23140-79 and AAR23722.

XX Sequence 545 AA;

Query Match 96.0%; Score 48; DB 13; Length 545;
 Best Local Similarity 90.9%; Pred. No. 0.45;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSXELSIPEE 11
 |||||||
 Db 390 LSQELSIPEE 400

RESULT 6
 AAR23156
 ID AAR23156 standard; Protein; 632 AA.
 XX

AC AAR23156;

DT 22-OCT-1992 (first entry)

DE Mutant thermostable DNA polymerase enzyme MET-THR 200 Tps17.

KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.

OS Thermus species SPS17.

Key Location/Qualifiers

FT Misc-d 'e'ence 1..2 /note= "residues 2-199 deleted from the native
 sequence"

PN WO9206200-A.

PD 16-APR-1992.

PF 30-SEP-1991; 91WO-US07035.

PR 28-SEP-1990; 90US-0590213.

PR 28-SEP-1990; 90US-0590466.

PR 28-SEP-1990; 90US-0590490.

PA (CETU) CETUS CORP.

PI Abramson RD, Gelfand DH;

DR WPI: 1992-150885/18.

DR N-PSDB; AAQ24009.

PS Claim 11; Page 59; 185pp; English.

XX The sequence is that of a mutant of Thermus species sps17 polymerase

CC mutant MET-THR 200 Tps17, having a different amt. of 5'-3' exonuclease

CC activity than the native enzyme. Thermostable DNA polymerases

CC are useful in many recombinant DNA techniques, esp. nucleic acid

CC amplification by PCR, self-sustained sequence replication (SSR)

CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity

CC may facilitate higher sensitivity allelic discrimination in a combined

CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of

CC 5'-3' exonuclease activity may be desirable in enzymes used in

CC homogeneous assays for the amplification and detection of a target

CC nucleic acid sequence. Mutation of the DNA encoding particular

CC regions of the enzymes can be used to prepare a range of recombinant

CC proteins having 5'-3' exonuclease activity to a complete lack of

CC activity.

CC See also AAR23140-79 and AAR23722.

XX Sequence 632 AA;

Query Match 96.0%; Score 48; DB 13; Length 632;

Best Local Similarity 90.9%; Pred. No. 0.53;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSXELSIPEE 11
 |||||||
 Db 478 LSQELSIPEE 488

RESULT 7
 AAR23155
 ID AAR23155 standard; Protein; 680 AA.
 XX

AC AAR23155;

DT 22-OCT-1992 (first entry)

DE Mutant thermostable DNA polymerase enzyme MET-LEU 152 Tps17.

KW 5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.

OS Thermus species SPS17.

Key Location/Qualifiers

FT Misc-difference 1..2 /note= "residues 2-151 deleted from the native
 sequence"

PN WO9206200-A.

PD 16-APR-1992.

PF 30-SEP-1991; 91WO-US07035.

PR 28-SEP-1990; 90US-0590213.

PR 28-SEP-1990; 90US-0590466.

PR 28-SEP-1990; 90US-0590490.

PA (CETU) CETUS CORP.

PI Abramson RD, Gelfand DH;

DR WPI: 1992-150885/18.

DR N-PSDB; AAQ24008.

PS Claim 11; Page 59; 185pp; English.

XX The sequence is that of a mutant of Thermus species sps17 polymerase

CC mutant MET-LEU 152 Tps17, having a different amt. of 5'-3' exonuclease

CC activity than the native enzyme. Thermostable DNA polymerases

CC are useful in many recombinant DNA techniques, esp. nucleic acid

CC amplification by PCR, self-sustained sequence replication (SSR)

CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity

CC may facilitate higher sensitivity allelic discrimination in a combined

CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of

CC 5'-3' exonuclease activity may be desirable in enzymes used in

CC homogeneous assays for the amplification and detection of a target

CC nucleic acid sequence. Mutation of the DNA encoding particular

CC regions of the enzymes can be used to prepare a range of recombinant

CC proteins having 5'-3' exonuclease activity to a complete lack of

CC activity.

CC See also AAR23140-79 and AAR23722.

XX Sequence 680 AA;

Query Match 96.0%; Score 48; DB 13; Length 680;

Best Local Similarity 90.9%; Pred. No. 0.58;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSXELSIPEE 11
 |||||||
 Db 526 LSQELSIPEE 536

RESULT 8
 AAR23154
 ID AAR23154 standard; Protein; 758 AA.

	Mutant thermostable DNA polymerase enzyme MEF-PHE 43 Tspst17.
XX	5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.
OS	Thermus species SPS17.
FH	Key location/Qualifiers
FT	Misc-difference 1..2 /note- "residues 2-43 deleted from the native sequence"
PJ	
PN	WO9206200-A.
PD	16-APR-1992.
PF	30-SEP-1991; 91WQ-US07035.
PR	28-SEP-1990; 90US-0590213.
PR	28-SEP-1990; 90US-0590466.
PR	28-SEP-1990; 90US-0590490.
PA	(CETU) CETUS CORP.
Pj	Abramson RD, Gelfand DH;
DR	WP1; 1992-150885/18.
DR	N-PSDB; AAQ24006.
PT	Thermostable DNA polymerases with altered 5'-3' exo nuclease activity - having conserved regions mutated or deleted, for use in e.g. PCR, sequencing and detection assays
xx	Claim 11: Page 59; 185pp; English.
CC	The sequence is that of a mutant of Thermus species sps17 polymerase mutant MER-PHE 43 Tspst17, having a different amt. of 5'-3' exonuclease activity than the native enzyme. Thermostable DNA polymerases are useful in many recombinant DNA techniques, esp. nucleic acid amplification by PCR, self-sustained sequence replication (SSR) and high temp. DNA sequencing. The absence of 5'-3' nuclease activity may facilitate higher sensitivity allelic discrimination in a combined polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of 5'-3' exonuclease activity may be desirable in enzymes used in homogeneous assays for the amplification and detection of a target nucleic acid sequence. Mutation of the DNA encoding particular regions of the enzymes can be used to prepare a range of recombinant proteins having 5'-3' exonuclease activity to a complete lack of activity.
CC	See also AAR23140-79 and AAR23722.
SQ	Sequence 788 AA;
Query Match	96.0%; Score 48; DB 13; Length 788;
Best Local Similarity	90.9%; Pred.No. 0.69;
Matches 10; Conservative	0; Mismatches 1; Indels 0; Caps 0;
Oy	1 LKXELSPYEE 11
Db	634 LSDELSPYEE 644
RESULT 10	
ID	AAR23152
AC	AAR23152 standard; Protein; 830 AA.
DT	22-OCT-1992 (first entry)
DE	Mutant thermostable DNA polymerase enzyme ASP43 Tspst17.
XW	5'-3'; exonuclease; PCR; amplification; SSR; sequencing; PLCR.

XX Thermus species sps17.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 37
 FT /note= "Gly in native sequence"
 XX
 PN W09206200-A.
 XX
 PD 16-APR-1992.
 XX
 PF 30-SEP-1991; 91WO-US07035.
 XX
 PR 28-SEP-1990; 90US-0590213.
 PR 28-SEP-1990; 90US-0590466.
 PR 28-SEP-1990; 90US-0590490.
 XX
 PA (CETU) CETUS CORP.
 XX
 PI Abramson RD, Gelland DH;
 XX
 DR WPI: 1992-150885/18.
 DR N-PSDB; AAQ24005.
 XX
 PT Thermostable DNA polymerases with altered 5'-3' exo nuclease
 PT activity - having conserved regions mutated or deleted, for use
 PT in e.g. PCR, sequencing and detection assays
 XX
 PS Claim 11; Page 59; 185pp; English.
 XX
 CC The sequence is that of a mutant of Thermus species sps17 polymerase
 CC mutant ASP43 Tps17, having a different amt. of 5'-3' exonuclease
 CC activity than the native enzyme. Thermostable DNA polymerases
 CC are useful in many recombinant DNA techniques, esp. nucleic acid
 CC amplification by PCR, self-sustained sequence replication (SSR)
 CC and high temp. DNA sequencing. The absence of 5'-3' nuclease activity
 CC may facilitate higher sensitivity allelic discrimination in a combined
 CC polymerase ligase chain reaction (PLCR) assay. An enhanced amt. of
 CC 5'-3' exonuclease activity may be desirable in enzymes used in
 CC homogeneous assays for the amplification and detection of a target
 CC nucleic acid sequence. Mutation of the DNA encoding particular
 CC regions of the enzymes can be used to prepare a range of recombinant
 CC proteins having 5'-3' exonuclease activity to a complete lack of
 CC activity.
 CC See also AAR23140-79 and AAR23722.
 CC
 XX
 SQ Sequence 830 AA;
 XX
 Query Match 96.0%; Score 48; DB 13; Length 830;
 Best Local Similarity 90.9%; Pred. No. 0.73;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 LSKELSIPIYEE 11
 II |||||
 DB 676 LSGELSIPIYEE 686
 XX
 RESULT 11
 AAR76060
 ID AAR76060 standard; Protein; 830 AA.
 XX
 AC AAR76060;
 XX
 DE 28-OCT-1995 (first entry)
 XX
 DE Tfil DNA polymerase.
 XX
 KW DNA-polymerase; Tfil; thermostable enzyme.
 XX
 OS Thermus filiformis.
 XX
 PN W09514770-A.
 XX

PD 01-JUN-1995.
 XX
 PE 23-NOV-1994; 94WO-NZ00135.
 XX
 PR 25-NOV-1993; 93NZ-0250288.
 XX
 PA (PACI-) PACIFIC ENZYMES 1993 LTD.
 XX
 PI Bergquist PL, Day DJ, Gibbs MD, Reeves RA, Saul DJ;
 XX
 DR WPI: 1995-206929/27.
 DR N-PSDB; AAQ92369.
 XX
 PT New heat-stable DNA polymerase from Thermus filiformis - has reverse
 PT transcriptase activity in the presence of magnesium ions.
 XX
 PS Claim 4; Fig 2; 40pp; English.
 XX
 CC Thermostable DNA-polymerase Tfil (DNA nucleotidyltransferase
 CC DNA-directed enzyme) has reverse-transcriptase activity in the
 CC presence of Mg2+. The enzyme is a versatile DNA-polymerase
 CC suitable for use in DNA synthesis from DNA and RNA templates,
 CC and for automated DNA sequencing. Both of these reactions are
 CC performed in the same buffer. The enzyme has a high salt
 CC tolerance and is also suitable for use in cycled fluorescent
 CC DNA sequencing using dye-primers and dye-terminators.
 CC
 XX
 SQ Sequence 830 AA;
 XX
 Query Match 96.0%; Score 48; DB 16; Length 830;
 Best Local Similarity 90.9%; Pred. No. 0.73;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 LSKELSIPIYEE 11
 II |||||
 DB 676 LSGELSIPIYEE 686
 XX
 RESULT 12
 AAR64273
 ID AAR64273 standard; Protein; 831 AA.
 XX
 AC AAR64273;
 XX
 DE 19-JUN-1995 (first entry)
 XX
 DE T. flavus DNA polymerase.
 XX
 KW DNA-polymerase; DNAE; Tfil; DNA cleavage; RNA cleavage; 5' nuclease.
 XX
 OS Thermus flavus AT-62.
 XX
 PN W09429482-A.
 XX
 PD 22-DEC-1994.
 XX
 PE 06-JUN-1994; 94WO-US06253.
 XX
 PR 04-JUN-1993; 93US-0073384.
 XX
 PA (THIR-) THIRD WAVE TECHNOLOGIES INC.
 XX
 PI Brow MAD, Dahlberg JE, Lyamichev VI;
 XX
 DR WPI: 1995-036504/05.
 DR N-PSDB; AAQ80750.
 XX
 PT 5' Nuclease(s) derived from thermostable DNA polymerase(s) - have
 PT cleavage activity with reduced synthetic ability, used for
 PT detection of specific target sequences.
 XX
 PS Disclosure; Page 83-86; 159pp; English.
 XX

CC The DNA and amino acid sequences of T. flavus AT-62 thermostable
 CC DNA-polymerase (DNAP) are given in AA080750 and AAB64273, respectively.
 CC 5' Nucleases, such as those given in AA080742-45, derived from
 CC thermostable DNAPs show reduced synthetic activity whilst retaining
 CC structure-specific cleavage activity.
 XX Sequence 831 AA;

Query Match 96.0%; Score 48; DB 16; Length 831;
 Best Local Similarity 90.9%; Pred. No. 0.73;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKELSIPEE 11
 || |||||
 DB 677 LSGELSIPEE 687

RESULT 13

AAW79961
 ID AAW79961 standard; Protein; 831 AA.

XX AAW79961;

XX 02-FEB-1999 (first entry)

XX Thermus flavus DNA polymerase.

XX Nucleic acid detection; multiple sequential invasive cleavage;
 KM DNA polymerase.

XX Thermus flavus.

XX WO9842873-A1.

XX 01-OCT-1998.

XX 24-MAR-1998; 98WO-US05809.

XX 24-MAR-1997; 97US-0823516.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Brow MAD, Hall JG, Kwiatkowski RM, Lyamichev VI;

XX Mast AL, Vavra SH;

XX WPI: 1998-557036/47.

XX N-PSDB: AAW65780.

XX Detecting target nucleic acid by sequence-specific cleavage of
 PT complex with two specific oligonucleotides - used to detect
 PT cytomegalovirus DNA

PS Example 2; Page 266-268; 524pp; English.

XX This is the amino acid sequence of the thermostable DNA polymerase
 CC (DNAP) of Thermus flavus (Tfl). The invention relates to means
 CC for the detection and characterisation of nucleic acid sequences,
 CC and variations in nucleic acid sequences. It also relates to
 CC methods for forming a nucleic acid cleavage structure on a target
 CC sequence and cleaving this structure in a site-specific manner,
 CC preferably using a thermostable structure-specific nuclease such as
 CC a modified Thermus DNAP that has reduced synthetic activity (see
 CC AAW65783-86). Cleavage of the cleavage structure by the nuclease
 CC indicates the presence of specific nucleic acid sequences or
 CC specific variants. The invention further relates to methods for
 CC the separation of nucleic acid molecules based on charge, methods
 CC for the detection of non-target cleavage products via the formation
 CC of a complete and activated protein binding region, and methods for
 CC the detection of nucleic acid from various viruses (e.g. human
 CC cytomegalovirus) in a sample. The method amplifies the detection
 CC molecule rather than the target itself, is less subject to
 CC contamination than exponential amplification processes, and allows
 CC many targets to be analysed in a single reaction.

XX Sequence 831 AA;

Query Match 96.0%; Score 48; DB 19; Length 831;
 Best Local Similarity 90.9%; Pred. No. 0.73;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSKELSIPEE 11
 || |||||
 DB 677 LSGELSIPEE 687

RESULT 14

AAW80428
 ID AAW80428 standard; Protein; 831 AA.

XX AAW80428;

XX 26-JAN-1999 (first entry)

XX DNA polymerase enzyme.

XX Thermus flavus DNA polymerase; nucleic acid detection;
 KM thermosable 5' nuclease.

XX Thermus flavus.

XX US5837450-A.

XX 17-NOV-1998.

XX 06-JUN-1995; 95US-0471066.

XX 06-JUN-1994; 94US-0254359.

XX 07-DEC-1992; 92US-0886330.

XX 04-JUN-1993; 93US-0073384.

XX 06-JUN-1995; 95US-0471066.

XX (THIR-) THIRD WAVE TECHNOLOGIES INC.

XX Brow MAD, Dahlborg JR, Lyamichev VI;

XX WPI: 1999-023438/02.

XX N-PSDB: AAW63401.

XX Detection of target nucleic acid molecules - uses modified
 PT thermostable enzymes with specific cleavage activity to create
 PT specific detection products from oligo-nucleotide(s) and target
 PT hybridisation

PS Disclosure; Fig 3; 91pp; English.

XX The present sequence represents a thermostable DNA polymerase enzyme.
 CC The sequence is used in the course of the invention. The specification
 CC describes a method for detecting the presence of a nucleic acid
 CC molecule. The method uses a thermostable 5' nuclease derived from a
 CC thermostable polymerase modified to have reduced synthetic activity,
 CC where the 5' nuclease is capable of cleaving a linear nucleic acid
 CC duplex structure to create a single, single-stranded cleavage product.
 CC The methods are used for the specific detection of nucleic acid
 CC sequences, via a cleavage-based procedure, but without the need for
 CC amplification of target sequences. Thermostable polymerases, altered to
 CC have nuclease, but not polymerase activity are preferably used due to
 CC their specificity. The cleavage product specifically formed is detected,
 CC preferably by the use of radioactively labelled oligonucleotides. These
 CC can be used in e.g. forensic testing or paternity determination.

XX Sequence 831 AA;

Query Match 96.0%; Score 48; DB 20; Length 831;
 Best Local Similarity 90.9%; Pred. No. 0.73;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSKELSIPEE 11
 || |||||
 Db 677 LSGELSIPEE 687

RESULT 15

AAK72875
 ID AAK72875 standard; Protein; 845 AA.

XX
 AC AAK72875;

XX
 DE 05-DEC-1995 (first entry)

XX
 DE Tpsp17 polymerase.

XX
 KW Polymerase; Thermus; enzyme; thermostable; exonuclease.

XX
 OS Thermus species sps17.

XX
 PN US5405774-A.

XX
 PD 11-APR-1995.

XX
 PF 22-AUG-1986; 86US-0899241.

XX
 PR 10-SEP-1993; 93US-0119754.

XX
 PR 22-AUG-1986; 86US-0899241.

XX
 PR 17-JUN-1987; 87US-0063509.

XX
 PR 12-JAN-1988; 88US-0143441.

XX
 PR 15-MAY-1990; 90US-0523394.

XX
 PR 28-SEP-1990; 90US-0590213.

XX
 PA (HOFF) HOFFMANN LA ROCHE INC.

XX
 PI Abramson RD, Gelfand DH, Greenfield IL;

XX
 DR WPI, 1995-154582/20.

XX
 DR N-PSDB; AAQ86927.

XX
 PT DNA for Thermus species sps17 DNA polymerase - modified to encode

XX
 PT an N-terminal deletion to reduce 5'-3' exonuclease activity

XX
 PS Disclosure; Columns 7-12; 26pp; English.

XX
 CC This sequence is the thermostable polymerase of Thermus species

XX
 CC sps17. The Gly residue at position 43 may be substituted for an

XX
 CC Asp residue. This substitution is shown to result in a 100 fold

XX
 CC decrease in the 5'-3' exonuclease activity with no change to the

XX
 CC polymerase activity of the enzyme.

XX
 SO Sequence 845 AA;

XX
 Query Match 96.0%; Score 48; DB 16; Length 845;

XX
 Best Local Similarity 90.9%; Pred. No. 0.74;

XX
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Caps 0;

XX
 OY 1 LSKELSIPEE 11

XX
 || |||||

XX
 Db 691 LSGELSIPEE 701

Search completed: January 15, 2003, 11:20:16
 Job time : 30.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:04:22 ; Search time 6.28571 seconds
(without alignments)
72.584 Million cell updates/sec

Title: US-09-823-649A-5
Perfect score: 49
Sequence: 1 LSVRLGXPKYKE 11

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	77.6	360	1 YAOX_RHISN	P55609 rhizobium s
2	36	73.5	477	1 ANCT_RAT	P01015 ratu
3	36	73.5	2144	1 BR28_HUMAN	O91583 homo sapien
4	34	69.4	334	1 YNMG_ECOLI	P76193 escherichia
5	34	69.4	356	1 RECA_THEMA	P36203 thetomologa
6	34	69.4	381	1 RM03_CAEEL	P49404 caenorhabd
7	34	69.4	557	1 GGAL_YEAST	O06336 saccharomyc
8	33	67.3	265	1 GYSH_RHISN	P55611 rhizobium s
9	33	67.3	621	1 YCXM_PORPU	P15224 porphyra pu
10	32	65.3	317	1 ARB_PIG	P18650 sus scrofa
11	32	65.3	341	1 HEAD_BPPH8	P05481 bacterioph
12	32	65.3	473	1 RNEC_BUCAL	P57215 buchnera ap
13	32	65.3	500	1 PC11_ARATH	P57681 arabidopsis
14	32	65.3	772	1 FROB_SALTY	P74884 salmonella
15	32	65.3	773	1 FROB_SALTY	P33650 escherichia
16	32	65.3	964	1 ION2_MAIZE	P93648 zea mays (m
17	32	65.3	1034	1 CAPP_SYNY3	P74299 synecocyst
18	32	65.3	1165	1 RPO2_SCHPO	Q10233 schizosach
19	32	65.3	1226	1 DSR4_HUMAN	P55265 homo sapien
20	31	63.3	123	1 GALA_BOVIN	P11242 bos tauris
21	31	63.3	123	1 GALA_PIG	P07480 sus scrofa
22	31	63.3	155	1 CHEM_CAUCR	O87715 caulobacter
23	31	63.3	249	1 APY1_PEA	P48534 pisum sativ
24	31	63.3	259	1 MORE_HELPJ	O92114 helicobacte
25	31	63.3	259	1 MORE_HELPJ	O29563 helicobacte
26	31	63.3	302	1 YCAN_ECOLI	P75836 escherichia
27	31	63.3	313	1 COBD_MYCTU	Q10518 mycobacteri
28	31	63.3	451	1 EX7L_NEIMB	O91Y22 neisseria m
29	31	63.3	511	1 P60_LISGR	O01835 listeria gr
30	31	63.3	520	1 VP7_MTV	P13092 wound tumor
31	31	63.3	759	1 PDR1_THENC	O91564 thermoplas
32	31	63.3	819	1 RNEC_HAELN	P71397 haemophilus
33	31	63.3	984	1 SRCA_AQUAE	O67718 aquifex aeo

34	30	61.2	237	1 HIS4_THERT	O9rp5 thermoaer
35	30	61.2	290	1 LIVE_ARCFU	O29329 archaeoglob
36	30	61.2	301	1 THTR_CORGL	P71121 coxynabate
37	30	61.2	340	1 TDDR_ECOLI	P77308 pseudomonas
38	30	61.2	374	1 CATB_PSEPU	P08310 pseudomonas
39	30	61.2	385	1 SUCC_AQUAE	O67546 aquifex aeo
40	30	61.2	390	1 EMRA_HAEIN	P44928 haemophilus
41	30	61.2	393	1 DCLQ_SELRU	O50657 scienomomas
42	30	61.2	406	1 FBM3_HUMAN	O9ukb7 homo sapien
43	30	61.2	559	1 DNL1_PYROO	O59288 pyrococcus
44	30	61.2	561	1 DNL1_PYRPU	P56709 pyrococcus
45	30	61.2	613	1 FIRP_ADEMI	P19721 mouse adeno

ALIGNMENTS

RESULT 1	YAOX_RHISN	STANDARD:	PRT:	360 AA.
ID	YAOX_RHISN			
AC	P55609:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Hypothetical 39.9 kDa protein YAOX.			
GN	YAOX.			
OS	Rhizobium sp. (strain NGR234).			
OG	Plasmid sym pNGR234a.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Rhizobiaceae; Rhizobium.			
OX	NCBI_TaxID=394;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97305956; PubMed=9163424;			
RA	Freiberg C.A., Fellay R., Balroch A., Broughton W.J., Rosenthal A.,			
RA	Perret X.;			
RT	"Molecular basis of symbiosis between Rhizobium and legumes.";			
RL	Nature 387:394-401(1997).			
CC	-1- FUNCTION: COULD BE A NAD-DEPENDENT OXIDOREDUCTASE.			
CC	-1- SIMILARITY: LOW, TO R. MELIOTI RHIZOPINE CATABOLISM PROTEIN MOCA			
CC	AND S. GRISPEUS STREPTOMYCIN BIOSYNTHESIS PROTEIN STR.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib.ch).			
CC				
DR	EMBL; AE000689; AAB91810.1; -			
DR	InterPro: IPR000683; GPO_IDH_MOCA.			
DR	InterPro: IPR004104; GPO_IDH_MOCA.C.			
DR	Pfam: PF01408; GPO_IDH_MOCA.1.			
DR	Pfam: PF02894; GPO_IDH_MOCA.C.1.			
KW	Hypothetical protein; Plasmid; Oxidoreductase; NAD.			
FT	NP_BIND 4 22 NAD (BY SIMILARITY).			
SQ	SEQUENCE 360 AA: 39853 MW: 59996F5082DE26CA CRC64;			
Query Match	77.6%; Score 38; DB 1; Length 360;			
Best Local Similarity	88.9%; Pred. No. 1.9;			
Matches	8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
DB	2 SVRLGXPKY 10			
QY				
DB	337 SVRLGXPKY 345			
RESULT 2	ANCT_RAT	STANDARD:	PRT:	477 AA.
ID	ANCT_RAT			
AC	P01015:			
DT	21-JUL-1986 (Rel. 01, Created)			

21-JUL-1986 (Rel. 01, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Angiotensinogen precursor [contains: Angiotensin I (Ang I);
 DE Angiotensin II (Ang II); Angiotensin III (Ang III) (Des-Asp11)-
 DE Angiotensin II(1)]
 GN ACT OR SERPIN8.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_Taxid=10116;
 RN SEQUENCE FROM N.A.
 RC STRAIN=Wistar.
 RX MEDLINE=83169849; PubMed=6572971;
 RA Okubo H., Kageyama R., Ujinara M., Hirose T., Inayama S.,
 RA Nakamishi S.;
 RT "Cloning and sequence analysis of cDNA for rat angiotensinogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2196-2200(1983).
 RN [2]
 RP SEQUENCE OF 25-34.
 RX MEDLINE=73060322; PubMed=4344907;
 RA Nakayama T., Nakajima T., Sokabe H.;
 RT "Comparative studies on angiotensins. II. Structure of rat
 RT angiotensin and its identification by DNS-method.";
 RL Chem. Pharm. Bull. 20:1579-1581(1972).
 CC -1- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
 CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
 CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
 CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
 CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
 CC BALANCE OF BODY FLUIDS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Synthesized by the liver and secreted in the
 CC plasma.
 CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: L00094; AAA98779.1; -
 DR EMBL: L00091; AAA98779.1; JOINED.
 DR EMBL: L00092; AAA98779.1; JOINED.
 DR EMBL: L00093; AAA98779.1; JOINED.
 DR PIR: A01251; ANRT.
 DR InterPro: IPR000227; Angiotensin.
 DR InterPro: IPR000215; Serpin.
 DR Pfam: PF00079; Serpin_1.
 DR PRINTS: PR00654; ANGIOTENSIN.
 DR SMART: SM00093; SERPIN_1.
 DR PROSITE: PS00284; SERPIN; FALSE NEG.
 KW Vasconstrictor; Glycoprotein; Plasma; Serpin; Signal.
 FT STGNAL. 1 24
 FT CHAIN 25 477 ANGIOTENSINOGEN.
 FT PEPTIDE 25 34 ANGIOTENSIN I.
 FT PEPTIDE 25 32 ANGIOTENSIN II.
 FT PEPTIDE 26 32 ANGIOTENSIN III.
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 477 AA; 51981 MW; 689051A578BD63D CRC64;

Query Match 73.5%; Score 36; DB 1; Length 477;
 Best Local Similarity 72.7%; Pred. No. 6.7;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 LSVRLGXPKYKE 11
 I : | | | | |
 DB 148 LQYLGLGVPKYKE 158

RESULT 3
 BP28_HUMAN STANDARD; PRT; 2144 AA.
 ID BP28_HUMAN
 AC 09H583; 09NM23;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein BAP28.
 GN BAP28.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS S-1694; A-1854; D-1967 AND G-2017.
 RA Bougueteloret I., Chumakov I., Barry C., Cohen-Akenine A.;
 RT "A novel BAP28 gene and protein.";
 RL Patent number WO0100669, 04-JAN-2001.
 RN [2]
 RP SEQUENCE OF 1534-2144 FROM N.A.
 RA Cobley V.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1777-2144 FROM N.A.
 RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagaetsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Yamamoto J., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
 RA Niimura K., Iwayanagi T.;
 RT "NEO human cDNA sequencing project.";
 RL Submitted (FEH-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE BAP28 FAMILY.
 CC
 CC -1- SIMILARITY: CONTAINS 1 HEAT REPEAT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AX067150; CAC26776.1; -
 DR EMBL: AL136105; CAC15948.1; -
 DR EMBL: AK001221; BA091564.1; ALT INIT.
 DR InterPro: IPR000357; HEAT_REPEAT.
 DR PROSITE: PS0077; HEAT_REPEAT; FALSE NEG.
 KW Polypeptidism.
 FT REPEAT 2106 2142 HEAT.
 FT VARIANT 1694 1694 N -> S.
 FT VARIANT 1854 1854 /FTID-VAR_010939.
 FT VARIANT 1967 1967 V -> A.
 FT VARIANT 1967 1967 N -> D.
 FT VARIANT 2017 2017 /FTID-VAR_010941.
 FT VARIANT 2017 2017 E -> G.
 SQ SEQUENCE 2144 AA; 242355 MW; D66816E78D8C9B7 CRC64;

Query Match 73.5%; Score 36; DB 1; Length 2144;
 Best Local Similarity 63.6%; Pred. No. 31;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 LSVRLGXPKYKE 11
 I : | | | | |
 DB 923 LFINLGSVPYKE 933

RESULT 4
 YNIG_ECOLI

```

ID  YNHG_ECOLI      STANDARD:      PRT:      334 AA.
AC  P76133; P76899; P76900;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein ynhg precursor.
GN  YNHG OR B1678.
OS  Escherichia coli.
OC  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC  Escherichia.
OX  NCBI_TaxID=562;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=K12 / MG1655;
RX  MEDLINE=97426617; PubMed=9278503;
RA  Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA  Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA  Gregor J., Davis N.M., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA  Mau B., Shao Y.;
RT  "The complete genome sequence of Escherichia coli K-12.";
RL  Science 277:1453-1474(1997).
RN  [2]
RP  SEQUENCE OF 1-233 FROM N.A.
RC  STRAIN=K12;
RX  MEDLINE=97251357; PubMed=9097039;
RA  Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA  Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA  Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA  Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA  Oshima T., Saito N., Sankel G., Seki Y., Sivasubraman S.,
RA  Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA  Yamamoto Y., Horiuchi T.;
RT  "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT  corresponding to the 28.0-40.1 min region on the linkage map.";
RL  DNA Res. 3:363-377(1996).
CC  -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
CC  -1- SIMILARITY: BELONGS TO THE EREK/YBIS/CFE/YNHG FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE000263; AAC74748.1; -
DR  EMBL; D90811; BAAL5458.1; -
DR  EMBL; D90812; BAAL5464.1; ALT_SEQ.
DR  ECGene; BE14015; ynhg.
DR  InterPro; IPR002482; LysM.
DR  Pfam; PF01476; LysM; 1.
DR  SMART; SM00257; LysM; 1.
DR  SMART; SM00257; LysM; 1.
KW  Hypothetical protein; Periplasmic; Signal; Complete proteome.
FT  SIGNAL      1      23      POTENTIAL.
FT  CHAIN       24      334      HYPOTHETICAL PROTEIN YNHG.
FT  SEQUENCE    334 AA; 36082 MW; 25078BBA293895F CRC64;
SQ
Query Match      69.4%; Score 34; DB 1; Length 334;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY  2 SVRLGXPK 10
DB  223 SVRLGXPK 231

```

```

DE  16-OCT-2001 (Rel. 40, Last annotation update)
DE  RecA protein (Recombinase A).
GN  RECA OR TM1859.
OS  Thermotoga maritima.
OC  Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC  Thermotogaceae; Thermotoga.
OX  NCBI_TaxID=2336;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95014407; PubMed=792298;
RA  Wetmur J.G., Wong D.M., Ortiz B., Tong J., Reichert F., Gelfand D.H.;
RT  "Cloning, sequencing, and expression of RecA proteins from three
RT  distantly related thermophilic eubacteria.";
RL  J. Biol. Chem. 269:25928-25935(1994).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MSB / DSM 3109;
RX  MEDLINE=99287316; PubMed=10360571;
RA  Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA  Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA  McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA  Stewart A.W., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA  Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA  Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT  "Evidence for lateral gene transfer between Archaea and Bacteria from
RT  genome sequence of Thermotoga maritima.";
RL  Nature 399:323-329(1999).
CC  -1- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF
CC  SINGLE-STRANDED DNA, THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED
CC  DNA BY DUPLEX DNA, AND THE ATP-DEPENDENT HYBRIDIZATION OF
CC  HOMODUPLEX SINGLE-STRANDED DNAs. IT INTERACTS WITH LEXA CAUSING
CC  ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE RECA FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; L23425; AAA27417.1; -
DR  EMBL; AE001823; AAD36921.1; -
DR  HSSP; P26345; 1G19.
DR  TIGR; TM1859; -
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001554; RecA.
DR  Pfam; PF00154; RecA; 1.
DR  PRINTS; PR00142; RECA.
DR  PRODom; PD000229; RECA; 1.
DR  SMART; SM00382; AAA; 1.
DR  PROSITE; PS00321; RECA_1; 1.
DR  PROSITE; PS00162; RECA_2; 1.
DR  PROSITE; PS0163; RECA_3; 1.
KW  DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding;
KW  Complete proteome.
FT  NP_BIND      68      75      ATP (BY SIMILARITY).
FT  SEQUENCE    356 AA; 38797 MW; A2B70853C26C95FD CRC64;
SQ
Query Match      69.4%; Score 34; DB 1; Length 356;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY  1 LSVRLGXPK 11
DB  225 MEYRGEPIKE 235

```

```

RESULT 5
ID  RECA_THEMA      STANDARD:      PRT:      356 AA.
AC  P36203;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)

```

```

RESULT 6
ID  RM03_CAEEL      STANDARD:      PRT:      381 AA.

```

AC P49404:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative mitochondrial 60S ribosomal protein L3.
 GN C26B6.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chordata; Rhabdillidae; Rhabdillidae;
 OC Rhabdillidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Fulton L.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE L3P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U13875; AAA21160.2; -
 CC DR WormPep: C26B6.6; CE26871.
 CC DR InterPro: IPR000597; Ribosomal_L3.
 CC DR Pfam: PF00297; Ribosomal_L3; 1.
 CC DR ProDom: PD001374; Ribosomal_L3; 1.
 CC DR ProSITE: PS00474; RIBOSOMAL_L3; FALSE_NEG.
 CC KW Ribosomal protein; Mitochondrion.
 CC SQ SEQUENCE 381 AA; 42288 MW; 3F57FBC8CF2B667 CRC64;
 Query Match 69.4%; Score 34; DB 1; Length 381;
 Best local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VRLGXPKE 11
 ||| ||||
 DB 164 VRLGXPKE 172
 RESULT 7
 ID GAL_YEAST STANDARD; PRT; 557 AA.
 AC 006336;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ADP-ribosylation factor binding protein GAL1 (golgi-localized, gamma
 DE ear-containing, Arp-binding protein 1).
 GN GAL1 OR YDR358W OR D9476.2.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
 RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
 RA Johnston L., Langston Y., Latreille P., Le T., Madis E., Menezes S.,
 RA Miller N., Nhan M., Pauley A., Peluso D., Rinken U., Riles L.,
 RA Tait A., Trevisan E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
 RA Wilson R., Waterston R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=2021637; PubMed=10747088;
 RA Hirst J., Lai W.M.Y., Bright N.A., Totty N., Seaman M.N.J.,
 RA Robinson M.S.;
 RT "A family of proteins with gamma-adaptin and VHS domains that
 RT facilitate trafficking between the trans-Golgi network and the
 RT vacuole/lysosome";
 RL J. Cell Biol. 149:67-80(2000).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=21064505; PubMed=11124697;
 RA Zhdankina O., Strand N.L., Redmond J.M., Boman A.L.;
 RT "Yeast GGA proteins interact with GTP-bound Arf and facilitate
 RT transport through the Golgi";
 RL Yeast 18:1-18(2001).
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF MEMBRANE TRAFFIC
 CC THROUGH THE TRANS-GOLGI NETWORK.
 CC -1- SUBUNIT: BINDS TO ARF1 AND ARF2.
 CC -1- SIMILARITY: CONTAINS 1 GAMMA-ADAPLIN C-TERMINAL DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 VHS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U28372; AAB64793.1; -
 CC DR SGD: S0002766; GAL1.
 CC DR InterPro: IPR004152; GAT dom.
 CC DR InterPro: IPR001121; Gamma_adaptin_C.
 CC DR InterPro: IPR002014; HRS.
 CC DR Pfam: PF00790; VHS; 1.
 CC DR Pfam: PF02883; Alpha_adaptin_C2; 1.
 CC DR Pfam: PF03127; GAT; 1.
 CC DR ProDom: PD003686; HRS; 1.
 CC DR ProDom: PD021457; Gamma_adaptin_C; 1.
 CC DR SMART: SM00288; VHS; 1.
 CC DR PROSITE: PS0179; VHS; 1.
 CC KW Protein transport; golgi stack; Coiled coil.
 CC FT DOMAIN 29 165 VHS.
 CC FT DOMAIN 185 205 COILED COIL (POTENTIAL).
 CC FT DOMAIN 231 259 COILED COIL (POTENTIAL).
 CC FT DOMAIN 442 550 GAMMA-ADAPLIN C-TERMINAL.
 CC SQ SEQUENCE 557 AA; 62319 MW; DBA76787998BDB7 CRC64;
 Query Match 69.4%; Score 34; DB 1; Length 557;
 Best local Similarity 60.0%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSVLGGXPVK 10
 ||| ||| |||
 DB 523 LSVLGGXPVK 532
 RESULT 8
 ID OTSB_RHISN STANDARD; PRT; 265 AA.
 AC P55611;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Probable trehalose-phosphatase (EC 3.1.3.12) (Trehalose 6-phosphate
 DE phosphatase) (TPP).
 GN OTSB OR Y4PB.
 OS Rhizobium sp. (strain NGR234).
 OG Plasmid sym pNCR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Felly R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- CATALYTIC ACTIVITY: Trehalose 6-phosphate + H(2)O = trehalose +
CC phosphate.
CC -1- SIMILARITY: TO THE E. COLI (OTS) AND YEAST ENZYME (TPS2).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000090; AAB91812.1;
DR InterPro: IPR003337; Trehalose_Phase.
DR Pfam: PF02358; Trehalose_Phase; 1.
DR TrEMBL: TIGR00685; otsb; 1.
KW Hydrolyase; Plasmid.
SQ SEQUENCE 265 AA; 27957 MW; AD21CF43EF7D794A CRC64;

Query Match 67.3%; Score 33; DB 1; Length 265;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPV 9
Db 230 LSVRLGXPV 238 --
||:|:|
||:|:|

RESULT 9
YCNB_PORPU STANDARD; PRT; 621 AA.
AC P51234;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 73.8 kDa protein in YCN10-PSBI intergenic region
DE (ORF621).
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RA Reich M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U38804; AAC08120.1;
DR Hypothetical protein; Chloroplast.
SQ SEQUENCE 621 AA; 73754 MW; 36BAF5EBE64F4BFC CRC64;

Query Match 67.3%; Score 33; DB 1; Length 621;
Best Local Similarity 55.6%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPV 9

```

```

Db 111 LSVRLGXPV 119
||:|:|
||:|:|

RESULT 10
APE_PIG STANDARD; PRT; 317 AA.
AC P18650; O19099;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Apolipoprotein E precursor (Apo-E).
GN APOE.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Brzowska A.M., Grimholt U., Kulseth M.A., Wold I., Rogne S.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=98347372; PubMed=9682450;
RA Ramsondar J.J., Rucker E.B., Vasquez J.C., Gallagher D.S.,
RA Grimm D.R., Lunney J.K., Schlook L.B., Piedrahita J.A.;
RT "Isolation and genetic characterization of the porcine apolipoprotein
RT E gene.";
RL AnIm. Genet. 29:43-47(1998).
RN [3]
RP PRELIMINARY SEQUENCE OF 19-31.
RX MEDLINE=81021043; PubMed=7417263;
RA Weisgraber K.H., Troxler R.F., Rall S.C., May R.W.;
RT "Comparison of the human, canine and swine E apoproteins.";
RL Biochem. Biophys. Res. Commun. 95:374-380(1980).
CC -1- FUNCTION: APO-E MEDIATES BINDING, INTERNALIZATION, AND CATABOLISM
CC OF LIPOPROTEIN PARTICLES. IT CAN SERVE AS A LIGAND FOR THE LDL(APO
CC B/E) RECEPTOR AND FOR THE SPECIFIC APO-E RECEPTOR (CHYLOMICRON
CC REMNANT) OF HEPATIC TISSUES.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SECRETED IN PLASMA.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X72835; CAA51356.1;
DR EMBL: U70240; AAC29512.1;
DR PIR: A05312; A05312.
DR PIR: S33450; S33450.
DR HSSP: P02649; 10EF.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
KW Glycoprotein; Plasma; Lipid transport; HDL; VLDL; Chylomicron;
KW Hepatic-binding; Repeat; Signal.
FT SIGNAL 1..18
FT CHAIN 19..317
FT DOMAIN 157..167
FT DOMAIN 161..164
FT DOMAIN 228..235
FT DOMAIN 79..254
FT REPEAT 79..100
FT REPEAT 101..122
FT REPEAT 123..144
FT REPEAT 145..166
FT REPEAT 167..188
FT REPEAT 189..210

```

RN	[1]
RP	SEQUENCE FROM N.A.

RA	Tsolis R., Baumber A.J., Heffron F., Stojiljkovic I.;
CC	"Contribution of rona- and feo-mediated iron uptake to growth of
FT	<i>Salmonella typhimurium</i> in the mouse.";
RL	Infect. Immun. 64:4549-4556(1996).
CC	-1- FUNCTION: PROBABLE GTP-DRIVEN TRANSPORTER OF FERROUS ION.
CC	-1- SUBCELLULAR LOCATION: integral membrane protein. inner membrane.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-sib.ch).
EMBL	AE008861; AAL22368.1; .
DR	EMBL; U70434; BAB09706.1; .
DR	StyGene; SC10640; Feob.
DR	InterPro; IPR003373; Feob.
DR	InterPro; IPR005289; GTP-binding_dom.
DR	Pfam; PF02421; Feob; 2. Small-GTP.
DR	TIGRPFAMS; TIGR00231; small_gtp_1.
DR	TIGRPFAMS; TIGR00437; feob; 1.
KW	Iron transport; transport; transmembrane; inner membrane; GTP-binding;
KW	Complete proteome.
FT	NP_BIND 10 17 GTP (POTENTIAL).
FT	NP_BIND 56 60 GTP (POTENTIAL).
FT	TRANSMEM 102 122 POTENTIAL.
FT	TRANSMEM 282 302 POTENTIAL.
FT	TRANSMEM 309 329 POTENTIAL.
FT	TRANSMEM 334 354 POTENTIAL.
FT	TRANSMEM 383 403 POTENTIAL.
FT	TRANSMEM 427 447 POTENTIAL.
FT	TRANSMEM 453 473 POTENTIAL.
FT	TRANSMEM 482 502 POTENTIAL.
FT	TRANSMEM 518 538 POTENTIAL.
FT	TRANSMEM 568 588 POTENTIAL.
FT	TRANSMEM 664 684 POTENTIAL.
FT	TRANSMEM 691 711 POTENTIAL.
FT	TRANSMEM 722 742 POTENTIAL.
FT	CONFLICT 33 36 WAGV -> NSAF (IN REF. 2).
FT	CONFLICT 97 97 N -> Y (IN REF. 2).
FT	CONFLICT 302 302 I -> V (IN REF. 2).
FT	CONFLICT 311 311 D -> N (IN REF. 2).
FT	CONFLICT 327 330 YLIH -> KPSN (IN REF. 2).
SQ	SHOHNCK 772 AA; 84200 MW; 94AHAA80CAA485C7 CRC64;
<hr/>	
Query Match 65 38; Score 32; DH 1; length 772;	
Best Local Similarity 77.88; Ref. No. 75;	
Matches 7; Conservative 0; Mismatches 2; Indels 0; gaps 0	
Dq	1 LSVRIOLXPY 9
Dd	137 LSTRLOCPY 145
<hr/>	
RESULT 15	
ID	FEOB_ECOLI STANDARD; PRT; 773 AA.
AC	P33650;
DT	01-FEB-1994 (Rel. 28, Created)
DI	01-FEB-1994 (Rel. 28, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Ferrous Iron Transport Protein B.
GN	Feob OR B3409.
OS	<i>Escherichia coli</i> .
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC	<i>Escherichia</i> .
CX	NCBL_TaxID=562;
LN	[1]
RP	SEQUENCE FROM N.A.

```

RC STRAIN-K12;
RX MEDLINE-94012482; PubMed-8407793;
RA Kammer M., Schoen C., Hantke K.;
RT "Characterization of the ferrous iron uptake system of Escherichia
  coli."
RL J. Bacteriol. 175:6212-6219(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
MEDLINE-97426617; PubMed-9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
  Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
  Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
  Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: PROBABLE GTP-DRIVEN TRANSPORTER OF FERROUS ION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
-----
DR EMBL: X71063; CAA50387.1; -.
DR EMBL: U18997; AAB58207.1; -.
DR EMBL: AF000416; AAC76434.1; -.
DR PIR: A36932; A36932.
DR Ecogene; EG12102; feob.
DR InterPro; IPR003373; Feob.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF02421; Feob; 1.
DR TIGRFAMS; TIGR00231; small_GTP; 1.
DR TIGRFAMS; TIGR00437; feob; 1.
DR TIGRFAMS; TIGR00650; MG442; 1.
KW Iron transport; Transport; Transmembrane; Inner membrane; GTP-binding;
KW Complete proteome.
FT NP_BIND 10 17 GTP (POTENTIAL).
FT NP_BIND 56 60 GTP (POTENTIAL).
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 282 302 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
FT TRANSMEM 334 354 POTENTIAL.
FT TRANSMEM 383 403 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 482 502 POTENTIAL.
FT TRANSMEM 518 538 POTENTIAL.
FT TRANSMEM 568 588 POTENTIAL.
FT TRANSMEM 664 684 POTENTIAL.
FT TRANSMEM 691 711 POTENTIAL.
FT TRANSMEM 722 742 POTENTIAL.
SQ SEQUENCE 773 AA; 84473 MW; 1150A14B131AB9AE CRC64;

```

```

Query Match 65.3%; Score 32; DB 1; Length 773;
Best local Similarity 77.8%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 LSVRLGXPV 9
Db 137 LSRAGCPV 145

```

Search completed: January 15, 2003, 11:21:28
Job time : 11.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 15, 2003, 10:29:39 ; Search time 25.5714 Seconds

(without alignments)
88,635 Million cell updates/sec

Title: US-09-823-649a-5

Perfect score: 49

Sequence: 1 LSVRLGXPYKE 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.potent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	95.9	09X1V4	09X1V4 thermotoga
2	38	77.6	09X138	09X138 neisseria m
3	38	77.6	09J5U2	09J5U2 neisseria m
4	37	75.5	08Y8V0	08Y8V0 brucella me
5	37	75.5	09WY86	09WY86 thermotoga
6	36	73.5	09H044	09H044 homo sapien
7	36	73.5	09Z673	09Z673 homo sapien
8	36	73.5	09A118	09A118 agrobacteri
9	36	73.5	09Y662	09Y662 aeropyrum p
10	35	71.4	051771	051771 borrelia bu
11	34	69.4	09W4R4	09W4R4 grosophila
12	34	69.4	09R8R8	09R8R8 alcaligenes
13	34	69.4	09R1V3	09R1V3 homo sapien
14	34	69.4	09R1N2	09R1N2 deinococcus
15	34	69.4	08R8H7	08R8H7 thermococcus
16	34	69.4	08X604	08X604 escherichia

17	34	69.4	362	16	0981G4	0981G4 rhizobium 1
18	34	69.4	481	16	0983E2	0983E2 rhizobium 1
19	34	69.4	676	16	097NCO	097NCO streptococ
20	34	69.4	985	10	094JAI	094JAI oryza sativ
21	34	69.4	997	4	015111	015111 homo sapien
22	34	69.4	1096	11	062688	062688 rattus norv
23	34	69.4	1208	4	075162	075162 homo sapien
24	34	69.4	1251	16	09PEA2	09PEA2 xylicella fas
25	33	67.3	244	5	08MOY9	08MOY9 ralsionia s
26	33	67.3	324	16	08YOV5	08YOV5 aprocalis
27	33	67.3	354	16	092J00	092J00 rhizobium m
28	33	67.3	380	17	09UYA3	09UYA3 pyrococcus
29	33	67.3	386	2	045308	045308 bacillus me
30	33	67.3	387	2	09K110	09K110 bacillus ce
31	33	67.3	456	13	091854	091854 xenopus lae
32	33	67.3	1490	13	079922	079922 xenopus lae
33	32	65.3	99	15	071221	071221 human immun
34	32	65.3	183	10	098RT1	098RT1 guilardia
35	32	65.3	225	2	09X556	09X556 enterococcu
36	32	65.3	249	16	09CL83	09CL83 pasteurilla
37	32	65.3	265	10	094BB9	094BB9 plitilus cr
38	32	65.3	305	2	08RL00	08RL00 providencia
39	32	65.3	311	16	08YBY5	08YBY5 brucella me
40	32	65.3	347	15	09YLO1	09YLO1 human immun
41	32	65.3	355	10	09RPH3	09RPH3 arabidopsis
42	32	65.3	363	16	09RS42	09RS42 deinococcus
43	32	65.3	370	10	09H835	09H835 arabidopsis
44	32	65.3	371	2	08RLH6	08RLH6 shewanella
45	32	65.3	429	16	08UG88	08UG88 agrobacteri

ALIGNMENTS

RESULT 1
ID 09X1V4 PRELIMINARY; PRT; 893 AA.
AC 09X1V4;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE DNA-directed DNA polymerase I.
GN TM619.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID-2336;
RP [1]
RC SBDUENCR PRM N.A.
RC STRAIN-MS8 / DSM 3109;
RX MBLINK-9928/316; PubMed-10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hall D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima."
RT Nature 393:323-329 (1999).
RL Nature 393:323-329 (1999).
EMBL: AF001805; AAD36686.1;
HSSP: P00582; IKLN.
TIGR: TM619;
DR InterPro: IPR002562; 3_5-exonuclease.
DR InterPro: IPR002421; 5_3-exonuclease.
DR InterPro: IPR001098; DNA pol.
DR InterPro: IPR002298; DNA pol.
DR InterPro: IPR000513; Exo_N.I.
DR InterPro: IPR003584; HNH_2.
DR Pfam: PF01612; 3_5-exonuclease; 1.
DR Pfam: PF01367; 5_3-exonuclease; 1.
DR Pfam: PF02739; 5_3-exonuclease; 1.
DR Pfam: PF00476; DNA pol. A; 1.
DR PRINTS: PR00868; DNAPOLI.

DR SMART: SM00474; 35EXOC; 1.
 DR SMART: SM00475; 53EXOC; 1.
 DR SMART: SM00279; HBH2; 1.
 DR SMART: SM00482; POLAC; 1.
 DR TIGRFAWS: TIGR00593; POLA; 1.
 DR TIGRFAWS: TIGR00393; Complete proteome.
 KW DNA-directed DNA polymerase.
 SQ SEQUENCE 893 AA; 102160 MW; 669EE226DBB39F3 CRC64;

Query Match 95.9%; Score 47; DB 16; Length 893;
 Best Local Similarity 90.9%; Pred. No. 0.35;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGXPKVE 11
 ||||| ||||
 Db 741 LSVRLGXPKVE 751

RESULT 2

O9K138 PRELIMINARY; PRT; 324 AA.
 AC O9K138:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Sugar isomerase, Kpsf/GutQ family.
 GN NMB0352.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

NCBI_TaxID=491;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tetteijn H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Pedon J.F., Dodson R.J.,
 RA Nelson W.C., Gillin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Uitterlinden T.R., Khouri H., Qin H., Yamathavan J.,
 RA Gill J., Scarlato V., Mastignani V., Piza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58".
 RL Science, 287:1809-1815(2000).
 DR EMBL: AE002392; AAF40795.1; -.
 DR TIGR: NMB0352; -.
 DR InterPro: IPR000644; CBS_domain.
 DR InterPro: IPR004800; KpsF.
 DR InterPro: IPR001347; SIS.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF01380; SIS; 1.
 DR SMART: SM00116; CBS; 2.
 DR TIGRFAWS: TIGR00393; KpsF; 1.
 KW Isomerase; Complete proteome.
 SQ SEQUENCE 324 AA; 34155 MW; 03A3A3BD83B553F5 CRC64;

Query Match 77.6%; Score 38; DB 16; Length 324;
 Best Local Similarity 70.0%; Pred. No. 9.8;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 SVRLGXPKVE 11
 :|||||:|
 Db 213 AVRGLTPKVE 222

RESULT 3

O9JSU2 PRELIMINARY; PRT; 324 AA.
 AC O9JSU2:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein NMA2135.

GN NMA2135.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NCBI_TaxID=65699;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-22491 / SEROTYPE 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
 RA Jagers K., Leathers S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrell B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RT meningitidis 22491".
 RL Nature 404:502-506(2000).
 DR EMBL: AL162758; CAB85347.1; -.
 DR InterPro: IPR000644; CBS_domain.
 DR InterPro: IPR001347; SIS.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF01380; SIS; 1.
 DR SMART: SM00116; CBS; 2.
 DR TIGRFAWS: TIGR00393; KpsF; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 324 AA; 34144 MW; D5989FBC97FF4876 CRC64;

Query Match 77.6%; Score 38; DB 16; Length 324;
 Best Local Similarity 70.0%; Pred. No. 9.8;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 SVRLGXPKVE 11
 :|||||:|
 Db 213 AVRGLTPKVE 222

RESULT 4

O8YBVO PRELIMINARY; PRT; 441 AA.
 AC O8YBVO:
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE NADH dehydrogenase (EC 1.6.99.3).
 GN BME10786.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=24459;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / AVCC 23456 / BIOTYPE 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Renik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letessou J.-J.,
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009713; AAL54028.1; -.
 DR InterPro: IPR001377; FAD_Pyl_redox.
 DR Pfam: PF00070; Pyl_redox; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 441 AA; 48477 MW; E9D846904C6CB5D7 CRC64;

Query Match 75.5%; Score 37; DB 16; Length 441;
 Best Local Similarity 63.6%; Pred. No. 22;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSVRLGXPKVE 11

Db 247 VERRUGTPVKD 257

RESULT 5

Q9WY86

PRELIMINARY;

PRT: 451 AA.

AC Q9WY86;

01-NOV-1999 (TREMblrel. 12, Created)

DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)

01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE Election transport complex protein, putative.

TM0244.

OS Thermotoga maritima.

Bacteria; Thermotogales; Thermotogaceae; Thermotoga.

OX NCBI_TaxID=2336;

RN [1]

SEQUENCE FROM N.A.

RC STRAIT-MSB8 / DSM 3109;

MEDLINE=99287316; PubMed=10360571;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,

Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,

RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrelt M.M.,

Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,

RA Heideberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,

Salberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

RT "Evidence for lateral gene transfer between Archaea and Bacteria from

genome sequence of Thermotoga maritima."

RL Nature 398:323-329(1999).

EMBL: AE001708; AAD3535.1; -

DR TIGR: TM0244;

InterPro: IPR001450; 4Fe4S-ferredoxin.

DR InterPro: IPR001949; Complex1_51K.

Pfam: PF001512; Complex1_51K; 1.

DR Pfam: PF00037; fer1.1.

PROSITE: PS00198; 4Fe4S-FERREDOXIN; 2.

KW Iron-sulfur; Complete proteome.

SEQUENCE 451 AA; 49515 MW; 752D7272CFC6E18 CRC64;

Query Match

Best Local Similarity 75.5%; Score 37; DB 16; Length 451;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVKE 11

313 LTVRIGTPVKD 323

RESULT 6

Q9H044

PRELIMINARY;

PRT: 277 AA.

AC Q9H044;

01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

01-MAR-2001 (TREMblrel. 16, Last annotation update)

DE Hypothetical 32.0 kDa protein (Fragment).

GN DKEZF762P093.

OS Homo sapiens (Human).

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE-MELANOMA;

Koerner K., Beyer A., Mewes H.W., Weil B., Wiemann S.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

EMBL: AL512733; CAC21665.1; -

KW Hypothetical protein.

NON_TER

FT SEQUENCE 277 AA; 31950 MW; 778CCFFC20D1A2E3 CRC64;

Query Match 73.5%; Score 36; DB 4; Length 277;

Best Local Similarity 63.6%; Pred. No. 22;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVKE 11

Db 37 LTVRIGTPVKD 47

RESULT 7

P82673

PRELIMINARY;

PRT: 323 AA.

AC P82673;

01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

01-OCT-2000 (TREMblrel. 15, Last annotation update)

DE Mitochondrial 28S ribosomal protein S28 (MRP-S28).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

SEQUENCE OF 1-23 FROM N.A.

Wei Y., Tsang Y.T.M., Wei G., Ku J.M., Ali-Osman F.R. Jr., Mazny D.,

RA Bouck J., Gibbs R.A., Margolis J.F.;

Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

SEQUENCE OF 8-212 FROM N.A.

RA NCJ/NINDS-CCAP;

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

SEQUENCE OF 183-305 FROM N.A.

RA MEDLINE=96026280; PubMed=7566098;

Adams M.D., Kerlavage A.R., Fleischmann R.D., Feldner R.A., Bult C.J.,

RA Lee N., Kirkness E.F., Weinstock K.G., Gocayne J.D., White O.,

Sutton G., Blake J.A., Brandon R.C., Chiu M.-W., Clayton R.A.,

RA Cline R.T., Collin M.D., Earle-Hughes J., Fine L.D., Fitzgerald L.M.,

Fitzhugh W.M., Fritchman J.L., Geoghegan N.S.M., Glodek A.,

RA Gnehm C.L., Hanna M.C., Hedblom E., Hinkley P.S., Kelley J.M.,

Klimke K.M., Kelly J.C., Liu L.-I., Marmaro S.M., Merrick J.M.,

RA Moreno-Palauques R.F., McDonald L.A., Nguyen D.T., Pellegrino S.M.,

Phillips C.A., Ryder S.E., Scott J.L., Saudek D.M., Shirley R.,

RA Small K.V., Spriggs T.A., Utterback T.R., Weidman J.F., Li Y.,

Bednarik D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.-J.,

RA Dinko D., Peng P., Perle A., Fischer C., Hastings G.A., He W.-W.,

Hu J.-S., Greene J.M., Gruber J., Hudson P., Kim A., Kozak D.L.,

RA Kunsch C., Li H., Li H., Meissner P.S., Olsen H., Raymond L.,

Wei Y.-F., Wang J., Xu C., Yu G.-L., Rubin S.M., Dillon P.J.,

RA Fannon M.R., Kosen C.A., Hestelline W.A., Fields C., Fraser C.M.,

Venter J.C.;

KT "Initial assessment of human gene diversity and expression patterns

based upon 93 million basepairs of cDNA sequence."

RL Nature 377:3-174(1995).

RN [4]

SEQUENCE OF 221-323 FROM N.A.

RA Hillier L., Clark N., Dubugue T., Elliston K., Hawkins M., Holtan M.,

Hultman M., Kucaba Y., Le M., Lennon G., Maita M., Parsons J.,

RA Rikfin L., Rohlfing T., Soares M., Tan F., Trevisan E., Waterston R.,

Williamson A., Wohlmann P., Wilson R.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [5]

IDENTIFICATION.

Koc E.C., Burkhardt M., Blackburn K., Moseley A., Koc H.,

RA Spremulli L.L.;

"Identification of four proteins from the small subunit of the

RT mammalian mitochondrial ribosome using a proteomics approach."

Submitted (JUL-2000) to the SWISS-PROT data bank.

CC -1-SUBCELLULAR LOCATION: MITOCHONDRION.

EMBL: BE242647; -; NOT_ANNOTATED_CDS.

DR EMBL: A109645; -; NOT_ANNOTATED_CDS.

EMBL: AA314022; -; NOT_ANNOTATED_CDS.

DR EMBL: N24248; -; NOT_ANNOTATED_CDS.

Ribosomal protein; Mitochondrion.

KW POLY-ALA.

DOMAIN

FT SEQUENCE 323 AA; 36844 MW; B29F819H914F2B49 CRC64;

Query Match 73.5%; Score 36; DB 4; Length 323;
 Best Local Similarity 63.6%; Pred. No. 26;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 11
 |||:| |||:
 Db 83 LPVRMGYPVKK 93

RESULT 8

Q9AF18

PRELIMINARY;

PRT; 430 AA.

AC Q9AF18;

DT 01-JUN-2001 (TEMBLrel. 17, Created)

DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)

DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE Putative transport protein Mep.

GN Mep.

OS Agrobacterium tumefaciens.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=358;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21289077; PubMed=11395455;

RA Luo Z.O., Farrand S.K.;

RT "The Agrobacterium tumefaciens rnd Homolog Is Required for Trar-Mediated Quorum-Dependent Activation of Ti Plasmid tra Gene Expression."

RL J. Bacteriol. 183:3919-3930(2001).

DR EMBL; AY026066; AAK13016.1; -.

DR InterPro: IPR003662; sub-Transporter.

DR Pfam: PF00083; sugar tr; 1.

SQ SEQUENCE 430 AA; 46176 MW; 313B2FE3D08A1E28 CRC64;

Query Match 73.5%; Score 36; DB 2; Length 430;

Best Local Similarity 63.6%; Pred. No. 35;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 11
 |||:| |||:
 Db 375 LQVRGYPVKE 385

RESULT 9

Q9YE62

PRELIMINARY;

PRT; 753 AA.

AC Q9YE62;

DT 01-NOV-1999 (TEMBLrel. 12, Created)

DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)

DT 01-OCT-2001 (TEMBLrel. 18, Last annotation update)

DE 753AA long hypothetical aldehyde oxidoreductase.

GN APO0708.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;

OC Desulfurococcaceae; Aeropyrum.

OX NCBI_TaxID=56636;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99310339; PubMed=10382966;

RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,

RA Jinno K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,

RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

RT "Complete genome sequence of an aerobic hyper-thermophilic

RT crenarchaeon, Aeropyrum pernix K1."

RL DNA Res. 6:83-101(1999).

DR EMBL; AP000060; BAA79684.1; -.

DR InterPro: IPR000674; Aldxan_dh_C.

DR Pfam: PF01315; Ald_Xan_dh_C; 1.
 DR Pfam: PF02738; Ald_Xan_dh_C2; 1.
 KN Complete proteome.
 SQ SEQUENCE 753 AA; 82419 MW; 6836618422F1322D CRC64;

Query Match 73.5%; Score 36; DB 17; Length 753;
 Best Local Similarity 70.0%; Pred. No. 62;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 10
 |||:| |||:
 Db 252 LAMRLGXPVKE 261

RESULT 10

O51771

PRELIMINARY;

PRT; 315 AA.

AC O51771;

DT 01-JUN-1998 (TEMBLrel. 06, Created)

DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE Xylose operon regulatory protein (XylR-2).

GN B0831.

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98065943; PubMed=9403685;

RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,

RA Duthie R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,

RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,

RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,

RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weisman J.,

RA Ueberlack T., Wetthey L., McDonald L., Artlich P., Bowman C.,

RA Smith H.O., Venter J.C.;

RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."

RT Nature 390:580-586(1997).

DR KMBL; AR001181; AAC67180.1; -.

DR TIGR; B0831; -.

DR InterPro: IPR000600; ROK_family.

DR Pfam: PF00480; ROK; 1.

KN Complete proteome.

SQ SEQUENCE 315 AA; 34455 MW; C132HA0030D54B75 CRC64;

Query Match 71.4%; Score 35; DB 16; Length 315;

Best Local Similarity 54.5%; Pred. No. 41;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 11
 |||:| |||:
 Db 190 VAMRLGXPVKE 200

RESULT 11

Q9W4R4

PRELIMINARY;

PRT; 110 AA.

AC Q9W4R4;

DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)

DE CG13021 protein (AT30881p).

GN CG13021.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Braconyera; Muscomorpha;

OX Ephyrididae; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolashakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokslein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu L., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Eranclisella C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,
 RA Jafali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Switz R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokslein P., Hong L., Adayani A., Carlson J.,
 RA Chapple M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Lao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pargasa V., Park S.,
 RA Patel S., Phouneavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.,
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003427; AAF45879.1; -;
 DR EMBL; AY075295; AAL68162.1; -;
 DR Flybase; FBgn0029669; CG13021.
 SQ SEQUENCE 110 AA; 11901 MW; B61518F244C49545 CRC64;

Query Match 69.4%; Score 34; DB 5; Length 110;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVK 10
 Db 81 INVRIGNPVK 90

RESULT 12
 O9RBF8 PRELIMINARY; PRT; 124 AA.
 AC O9RBF8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 13.7 kDa protein (Fragment).
 OS *Alcaligenes eutrophus* (Ralstonia eutropha).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.
 OX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-335;
 RA Seibert V., Schloemann M.;
 RT "Characterization of a Gene Cluster Encoding the Malylacetate
 RT Reductase from *Ralstonia eutropha* 335, an Enzyme Recruited for Growth
 RT with 4-Fluorobenzoate.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF130250; AAD55889.1; -;
 DR InterPro: IPR005119; IYSR_subst.
 DR Pfam: PF03466; IYSR_substrate; 1.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 124 AA; 13679 MW; 3DE99E10EF97EB8A CRC64;

Query Match 69.4%; Score 34; DB 2; Length 124;
 Best Local Similarity 87.5%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVK 8
 Db 11 INVRIGNPVK 18

RESULT 13
 O8TDV3 PRELIMINARY; PRT; 235 AA.
 AC O8TDV3;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative G-protein coupled receptor.
 GN GPCR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Takada S., Kadowaki S., Haga T., Takasu H., Mitaku S.;
 RT "Identification of G protein-coupled receptor genes from the human
 RT genome sequence.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB063587; BAB89300.1; -;
 DR K receptor.
 SQ SEQUENCE 235 AA; 24154 MW; 9M1071B0D706B308 CRC64;

Query Match 69.4%; Score 34; DB 4; Length 235;
 Best Local Similarity 77.8%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VRLGXPVK 11
 Db 180 VRLGXPVK 188

RESULT 14
 O9RTN2 PRELIMINARY; PRT; 277 AA.
 AC O9RTN2;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein DR1728.
 GN DR1728.
 OS *Deinococcus radiodurans*.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-R;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577 (1999).
 DR EMBL; AE002014; AAP11284.1; .
 DR TIGR; DR728; .
 DR InterPro; IPR005177; DUF299.
 DR Pfam; PF03618; DUF299; 1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 277 AA; 30170 MW; FB133B5339C6BF2B CRC64;

Query Match 69.4%; Score 34; DB 16; Length 277;
 Best Local Similarity 63.6%; Pred. No. 58;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 LSVRLGXPKVKE 11
 | ||| |||
 DB 242 LPARIGLPVKD 252

RESULT 15

Q8RBH7 PRELIMINARY; FR: 301 AA.
 AC Q8RBH7;
 DT 01-JUN-2002 (Tremblrel. 21. Created)
 DT 01-JUN-2002 (Tremblrel. 21. Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21. Last annotation update)
 DE Membrane proteins related to metalloendopeptidases.
 GN NLPD2 OR TTE0841.
 OS Thermotoga thermophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermotogales; Thermotogaceae; Thermotoga thermophilus.
 OX NCBI_TaxID=119072;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome.";
 RL Genome Res. 12:689-700(2002).
 DR EMBL; AE013050; AAM24098.1; .
 KW Complete proteome.
 SQ SEQUENCE 301 AA; 33068 MW; 080B87435B3197B CRC64;

Query Match 69.4%; Score 34; DB 16; Length 301;
 Best Local Similarity 60.0%; Pred. No. 63;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 LSVRLGXPKVKE 10
 : || : |||
 DB 205 ISVHTGTPVK 214

Search completed: January 15, 2003, 12:35:56
 Job time : 33.5714 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 16, 2003, 08:03:43 / Search time 1525.14 seconds

(without alignments)
209.902 Million cell updates/sec

Title: US-09-823-649a-5
Sequence: 49
1 LSVRLGXPYKE 11

Scoring table: BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-O=/usr2.1/USPIO.spool/US09823649/runat.14012003.151002.29139/app.query.fasta.1.1393
-DB=GenEmbl -OEMT=fastap -SUFFIX=rcg -MINMATCH=0.1 -LOOEXT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR=SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09823649 @ecgn.1.13378 @runat.14012003.151002.29139 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGCLOG -DEV_TIMEOUT=120
-WAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.cm:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.in:*
18: em.inu:*
19: em.jmu:*
20: em.lov:*
21: em.ov:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	length	DB	ID	Description
1	47	95.9	694	6	AR072499	AR072499 Sequence
2	47	95.9	1485	6	AR094392	AR094392 Sequence
3	47	95.9	1485	6	AR098606	AR098606 Sequence
4	47	95.9	1716	6	AR094391	AR094391 Sequence
5	47	95.9	1716	6	AR098605	AR098605 Sequence
6	47	95.9	1737	6	AR094373	AR094373 Sequence
7	47	95.9	1737	6	AR098587	AR098587 Sequence
8	47	95.9	1833	6	AR094371	AR094371 Sequence
9	47	95.9	1833	6	AR094377	AR094377 Sequence
10	47	95.9	1833	6	AR094379	AR094379 Sequence
11	47	95.9	1833	6	AR094382	AR094382 Sequence
12	47	95.9	1833	6	AR094384	AR094384 Sequence
13	47	95.9	1833	6	AR094386	AR094386 Sequence
14	47	95.9	1833	6	AR094389	AR094389 Sequence
15	47	95.9	1833	6	AR094390	AR094390 Sequence
16	47	95.9	1833	6	AR098585	AR098585 Sequence
17	47	95.9	1833	6	AR098591	AR098591 Sequence
18	47	95.9	1833	6	AR098593	AR098593 Sequence
19	47	95.9	1833	6	AR098596	AR098596 Sequence
20	47	95.9	1833	6	AR098598	AR098598 Sequence
21	47	95.9	1833	6	AR098600	AR098600 Sequence
22	47	95.9	1833	6	AR098603	AR098603 Sequence
23	47	95.9	1833	6	AR098604	AR098604 Sequence
24	47	95.9	1833	6	AR098620	AR098620 Sequence
25	47	95.9	2682	6	AR024935	AR024935 Sequence
26	47	95.9	2682	6	AR072494	AR072494 Sequence
27	47	95.9	2682	6	AR094366	AR094366 Sequence
28	47	95.9	2682	6	AR098580	AR098580 Sequence
29	47	95.9	2682	6	AR149833	AR149833 Sequence
30	47	95.9	2682	6	AR173616	AR173616 Sequence
31	47	95.9	2682	6	E31690	E31690 Denatured c
32	47	95.9	2682	6	E112096	E112096 Sequence 1
33	47	95.9	2682	6	E115436	E115436 Sequence 3
34	47	95.9	2682	6	141266	141266 Sequence 1
35	47	95.9	13495	1	AE001805	AE001805 Thermotog
36	44	89.8	110000	2	AC069419_0	AC069419 Homo sapi
37	44	89.8	110000	2	AC111232	AC111232 Homo sapi
38	44	89.8	164608	2	AC111233	AC111233 Homo sapi
39	44	89.8	190060	9	AC068258	AC068258 Homo sapi
40	44	89.8	296483	2	AC102022	AC102022 Mus muscu
41	44	87.8	152346	2	AC102022	AC102022 Mus muscu
42	42	85.7	152346	2	AB003544	AB003544 Oryza sat
43	41	83.7	135792	2	AB004731	AB004731 Oryza sat
44	41	83.7	158456	2	AB004731	AB004731 Oryza sat
45	41	83.7	165654	2	AC027747	AC027747 Homo sapi

RESULT 1

ALIGNMENTS

AR072499 AR072499 694 bp DNA linear PAT 28-AUG-2000
LOCUS AR072499 Sequence 17 from patent US 5948614.
ACCESSION AR072499
VERSION AR072499.1 GI:9999263
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 694)
AUTHORS Chatterjee,D.K.
TITLE Cloned DNA polymerases from thermotoga maritima and mutants thereof
JOURNAL Patent: US 5948614-A 17 07-SEP-1999;
FEATURES
source
1. 694
/organism="unknown"
BASE COUNT 229 a 136 c 179 g 150 t
ORIGIN
Alignment Scores:
Pred. No.: 1.12 Length: 694
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0
US-09-823-649a-5 (1-11) x AR072499 (1-694)
QY 1 LeuserValArgLeugly**ProVallysglu 11
|||||
Db 233 CTTCTGTGAGACTTGGAATACCGGTTAAAGAA 265
RESULT 2
LOCUS AR094392 1485 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 38 from patent US 6001645.
ACCESSION AR094392
VERSION AR094392.1 GI:10021267
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1485)
AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
TITLE Thermophilic DNA polymerases from thermotoga neapolitana
JOURNAL Patent: US 6001645-A 38 14-DEC-1999;
FEATURES
source
1. 1485
/organism="unknown"
BASE COUNT 490 a 312 c 365 g 318 t
ORIGIN
Alignment Scores:
Pred. No.: 2.36 Length: 1485
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0
US-09-823-649a-5 (1-11) x AR094392 (1-1485)
QY 1 LeuserValArgLeugly**ProVallysglu 11
|||||
Db 1024 CTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1056
RESULT 3
LOCUS AR098606 1485 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 38 from patent US 6077664.
ACCESSION AR098606
VERSION AR098606.1 GI:12808372

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1485)
AUTHORS Slater,M.R., Huang,F., Hartnett,J.R., Bolchakova,E., Storts,D.R.,
TITLE Thermophilic DNA polymerases from Thermotoga neapolitana
JOURNAL Patent: US 6077664-A 38 20-JUN-2000;
FEATURES
source
1. 1485
/organism="unknown"
BASE COUNT 490 a 312 c 365 g 318 t
ORIGIN
Alignment Scores:
Pred. No.: 2.36 Length: 1485
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0
US-09-823-649a-5 (1-11) x AR098606 (1-1485)
QY 1 LeuserValArgLeugly**ProVallysglu 11
|||||
Db 1024 CTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1056
RESULT 4
LOCUS AR094391 1716 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 36 from patent US 6001645.
ACCESSION AR094391
VERSION AR094391.1 GI:10021265
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1716)
AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
TITLE Thermophilic DNA polymerases from thermotoga neapolitana
JOURNAL Patent: US 6001645-A 36 14-DEC-1999;
FEATURES
source
1. 1716
/organism="unknown"
BASE COUNT 504 a 367 c 414 g 385 t
ORIGIN
Alignment Scores:
Pred. No.: 2.71 Length: 1716
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0
US-09-823-649a-5 (1-11) x AR094391 (1-1716)
QY 1 LeuserValArgLeugly**ProVallysglu 11
|||||
Db 1255 CTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1287
RESULT 5
LOCUS AR098605 1716 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 36 from patent US 6077664.
ACCESSION AR098605
VERSION AR098605.1 GI:12808371
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1716)
AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
TITLE Thermophilic DNA polymerases from thermotoga neapolitana
JOURNAL Patent: US 6077664-A 36 14-DEC-1999;
FEATURES
source
1. 1716
/organism="unknown"

REFERENCE 1 (bases 1 to 1716)
AUTHORS Slater,M.R., Huang,F., Hartnett,J.R., Bolchakova,E., Storts,D.R.,
Otto,P., Miller,K.M., Novikov,A. and Velikodvorskaya,G.A.
TITLE Thermophilic DNA polymerases from Thermotoga neapolitana
JOURNAL Patent: US 6077664-A 36 20-JUN-2000;
FEATURES Location/Qualifiers
SOURCE 1..1716
BASE COUNT 551 a 367 c 413 g 385 t
ORIGIN

Alignment Scores:
Pred. No.: 2.71 Length: 1716
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-5 (1-11) x AR098605 (1-1716)

QY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 1255 CTTCTGTGAGACTTGGAATACCGTTAAAGAA 1287

RESULT 6
LOCUS AR094373 1737 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 10 from patent US 6001645.
ACCESSION AR094373
VERSION AR094373.1 GI:10021231
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1737)
AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
TITLE Thermophilic DNA polymerases from thermotoga neapolitana
JOURNAL Patent: US 6001645-A 10 14-DEC-1999;
FEATURES Location/Qualifiers
SOURCE 1..1737
BASE COUNT 553 a 374 c 417 g 393 t
ORIGIN

Alignment Scores:
Pred. No.: 2.75 Length: 1737
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-5 (1-11) x AR094373 (1-1737)

QY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 1276 CTTCTGTGAGACTTGGAATACCGTTAAAGAA 1308

RESULT 7
LOCUS AR098587 1737 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 10 from patent US 6077664.
ACCESSION AR098587
VERSION AR098587.1 GI:12808353
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1737)
AUTHORS Slater,M.R., Huang,F., Hartnett,J.R., Bolchakova,E., Storts,D.R.,
Otto,P., Miller,K.M., Novikov,A. and Velikodvorskaya,G.A.
TITLE Thermophilic DNA polymerases from Thermotoga neapolitana

JOURNAL Patent: US 6077664-A 10 20-JUN-2000;
FEATURES Location/Qualifiers
SOURCE 1..1737
BASE COUNT 553 a 374 c 417 g 393 t
ORIGIN

Alignment Scores:
Pred. No.: 2.75 Length: 1737
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-5 (1-11) x AR098587 (1-1737)

QY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 1276 CTTCTGTGAGACTTGGAATACCGTTAAAGAA 1308

RESULT 8
LOCUS AR094371 1833 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 7 from patent US 6001645.
ACCESSION AR094371
VERSION AR094371.1 GI:10021227
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)
AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
TITLE Thermophilic DNA polymerases from thermotoga neapolitana
JOURNAL Patent: US 6001645-A 7 14-DEC-1999;
FEATURES Location/Qualifiers
SOURCE 1..1833
BASE COUNT 591 a 394 c 440 g 408 t
ORIGIN

Alignment Scores:
Pred. No.: 2.89 Length: 1833
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 6 Gaps: 0

US-09-823-649A-5 (1-11) x AR094371 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 1372 CTTCTGTGAGACTTGGAATACCGTTAAAGAA 1404

RESULT 9
LOCUS AR094377 1833 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 15 from patent US 6001645.
ACCESSION AR094377
VERSION AR094377.1 GI:10021239
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)
AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
TITLE Thermophilic DNA polymerases from thermotoga neapolitana
JOURNAL Patent: US 6001645-A 15 14-DEC-1999;
FEATURES Location/Qualifiers
SOURCE 1..1833
BASE COUNT 592 a 393 c 440 g 408 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094377 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

RESULT 10

LOCUS AR094379 1833 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 18 from patent US 6001645.

ACCESSION AR094379

VERSION AR094379.1 GI:10021241

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)

AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.

TITLE Thermophilic DNA polymerases from thermotoga neapolitana

JOURNAL Patent: US 6001645-A 18 14-DEC-1999;

FEATURES

source Location/Qualifiers

1..1833

BASE COUNT 590 a 395 c 440 g 408 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094379 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

RESULT 11

LOCUS AR094382 1833 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 22 from patent US 6001645.

ACCESSION AR094382

VERSION AR094382.1 GI:10021247

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)

AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.

TITLE Thermophilic DNA polymerases from thermotoga neapolitana

JOURNAL Patent: US 6001645-A 22 14-DEC-1999;

FEATURES

source Location/Qualifiers

1..1833

BASE COUNT 590 a 393 c 440 g 410 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094382 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

RESULT 12

LOCUS AR094384 1833 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 25 from patent US 6001645.

ACCESSION AR094384

VERSION AR094384.1 GI:10021251

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)

AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.

TITLE Thermophilic DNA polymerases from thermotoga neapolitana

JOURNAL Patent: US 6001645-A 25 14-DEC-1999;

FEATURES

source Location/Qualifiers

1..1833

BASE COUNT 591 a 394 c 439 g 409 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094384 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

RESULT 13

LOCUS AR094386 1833 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 26 from patent US 6001645.

ACCESSION AR094386

VERSION AR094386.1 GI:10021255

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)

AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.

TITLE Thermophilic DNA polymerases from thermotoga neapolitana

JOURNAL Patent: US 6001645-A 28 14-DEC-1999;

FEATURES

source Location/Qualifiers

1..1833

BASE COUNT 590 a 395 c 440 g 408 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094386 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094382 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

RESULT 12

LOCUS AR094384 1833 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 25 from patent US 6001645.

ACCESSION AR094384

VERSION AR094384.1 GI:10021251

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)

AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.

TITLE Thermophilic DNA polymerases from thermotoga neapolitana

JOURNAL Patent: US 6001645-A 25 14-DEC-1999;

FEATURES

source Location/Qualifiers

1..1833

BASE COUNT 591 a 394 c 439 g 409 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094384 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

RESULT 13

LOCUS AR094386 1833 bp DNA linear PAT 08-SEP-2000

DEFINITION Sequence 26 from patent US 6001645.

ACCESSION AR094386

VERSION AR094386.1 GI:10021255

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1833)

AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.

TITLE Thermophilic DNA polymerases from thermotoga neapolitana

JOURNAL Patent: US 6001645-A 28 14-DEC-1999;

FEATURES

source Location/Qualifiers

1..1833

BASE COUNT 590 a 395 c 440 g 408 t

ORIGIN

Alignment Scores:

Pred. No.: 2.89

Score: 47.00

Percent Similarity: 90.91%

Best Local Similarity: 90.91%

Query Match: 95.92%

DB: 6

US-09-823-649a-5 (1-11) x AR094386 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValLysGlu 11

DB 1372 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1404

US-09-823-649A-5 (1-11) x AR094386 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValIysGlu 11
 DB 1372 CTTCTGTGAGACTGTGGAATACCGGTAAAGAA 1404

RESULT 14

AR094389 1833 bp DNA linear PAT 08-SEP-2000
 LOCUS AR094389 Sequence 32 from patent US 6001645.
 DEFINITION AR094389
 ACCESSION AR094389
 VERSION AR094389.1 GI:10021261
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1833)
 AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
 TITLE Thermophilic DNA polymerases from thermotoga neapolitana
 JOURNAL Patent: US 6001645-A 32 14-DEC-1999;
 FEATURES location/Qualifiers
 source 1..1833
 /organism="unknown"

BASE COUNT 580 a 394 c 441 g 408 t
 ORIGIN

Alignment Scores:

Pred. No.: 2.89 Length: 1833
 Score: 47.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 95.92% Indels: 0
 DB: 6 Caps: 0

US-09-823-649A-5 (1-11) x AR094389 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValIysGlu 11
 DB 1372 CTTCTGTGAGACTGTGGAATACCGGTAAAGAA 1404

RESULT 15

AR094390 1833 bp DNA linear PAT 08-SEP-2000
 LOCUS AR094390 Sequence 34 from patent US 6001645.
 DEFINITION AR094390
 ACCESSION AR094390
 VERSION AR094390.1 GI:10021263
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1833)
 AUTHORS Slater,M.R., Huang,F. and Hartnett,J.R.
 TITLE Thermophilic DNA polymerases from thermotoga neapolitana
 JOURNAL Patent: US 6001645-A 34 14-DEC-1999;
 FEATURES location/Qualifiers
 source 1..1833
 /organism="unknown"

BASE COUNT 589 a 395 c 441 g 408 t
 ORIGIN

Alignment Scores:

Pred. No.: 2.89 Length: 1833
 Score: 47.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 95.92% Indels: 0
 DB: 6 Caps: 0

US-09-823-649A-5 (1-11) x AR094390 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValIysGlu 11
 DB 1372 CTTCTGTGAGACTGTGGAATACCGGTAAAGAA 1404

Search completed: January 16, 2003, 11:02:34
 Job time : 1532.14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2003, 11:21:44 [Search time 130.143 Seconds
(without alignments)]

190.345 Million cell updates/sec

Title: US-09-823-649A-5

Perfect score: 49

Sequence: 1 LSVRLGXPKVE 11

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: --
-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09823649/rnat.14012003.151001.29119/app.query.fasta.1.1393
-DB=N.Geneseq.101002 -QPM=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -DNITS=bits -STAR=1 -END=1 -MATH=1 -DLSUM=2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.ecgn.1.1.0.0@rnat.14012003.151001.29119 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -TARGET=NG -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-NARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

N.Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	47	95.9	694	17 AAT30049	The DNA polymerase
2	47	95.9	694	18 AAT60031	Thermotoga neapolitana
3	47	95.9	694	21 AAT88767	T. neapolitana The
4	47	95.9	1485	18 AAT58155	Thermotoga neapoli
5	47	95.9	1719	18 AAT58145	Thermotoga neapoli
6	47	95.9	1737	18 AAT58144	Thermotoga neapoli
7	47	95.9	1833	18 AAT58135	Thermotoga neapoli
8	47	95.9	1833	18 AAT58136	Thermotoga neapoli
9	47	95.9	1833	18 AAT58137	Thermotoga neapoli
10	47	95.9	1833	18 AAT58138	Thermotoga neapoli
11	47	95.9	1833	18 AAT58139	Thermotoga neapoli
12	47	95.9	1833	18 AAT58140	Thermotoga neapoli
13	47	95.9	1833	18 AAT58141	Thermotoga neapoli
14	47	95.9	1833	18 AAT58142	Thermotoga neapoli
15	47	95.9	1833	18 AAT58143	Thermotoga neapoli
16	47	95.9	1836	13 AAO24004	Mutant thermotabi
17	47	95.9	2466	13 AAO24002	Mutant thermotabi
18	47	95.9	2574	13 AAO24000	Mutant thermotabi
19	47	95.9	2628	13 AAO24001	Mutant thermotabi
20	47	95.9	2682	13 AAO22052	Sequence encoding
21	47	95.9	2682	13 AAO23999	Mutant thermotabi
22	47	95.9	2682	18 AAT60076	Thermotoga neapoli
23	47	95.9	2682	18 AAT58134	Thermotoga neapoli
24	47	95.9	2682	19 AAV52930	The DNA polymerase
25	47	95.9	2682	20 AAV06642	Thermotoga neapoli
26	47	95.9	2682	20 AAX16998	Thermotoga marlim
27	47	95.9	2682	21 AAZ88762	T. neapolitana The
28	47	95.9	2682	22 AAZ88762	Human immune/naema
29	39	79.6	23580	22 AAK83578	Human immune/naema
30	39	79.6	23580	22 AAK83578	Human immune/naema
31	38	77.6	609	21 AAZ82274	N. meningitidis pa
32	38	77.6	981	21 AAZ53405	Neisseria meningit
33	38	77.6	981	21 AAZ53406	Neisseria meningit
34	38	77.6	1269	21 AAC48606	Arabisopsis thalia
35	38	77.6	1286	21 AAC48617	Arabisopsis thalia
36	38	77.6	1856	21 AAC33590	Arabisopsis thalia
37	38	77.6	25509	21 AAZ81467	N. meningitidis pa
38	38	77.6	349980	21 AAV21607	Neisseria meningit
39	38	77.6	534720	19 AAV30458	Rhizobium species
40	38	77.6	536165	19 AAV30459	Rhizobium species
41	38	77.6	1437668	21 AAZ81490	N. meningitidis B
42	37	75.5	361	24 AAZ81490	Human ORX polynuc
43	37	75.5	696	24 AAZ81490	Helicillus citreus
44	37	75.5	1299	21 AAZ53365	Neisseria gonorrhoe
45	37	75.5	1534	21 AAZ53366	Neisseria meningit

ALIGNMENTS

RESULT 1
ID AAT30049 standard: DNA, 694 BP.

AC AAT30049:

DT 15-AUG-1996 (first entry)

DE The DNA polymerase C-terminal region DNA.

XX DNA polymerase; The; DNA sequencing; amplification;

KW thermostable enzyme; ss.

OS Thermotoga neapolitana strain DSM 5068.

XX Key Location/Qualifiers
FH 2..94
FT CDS /*tag= a

```

FT      /transl_except= (44..47, aa:Glu)
TT      /transl_except= (671..673, aa:Ser)
FT      /transl_except= (674..676, aa:11e)

XX      W0910640-A1.
PN      11-APR-1996.
XX      02-OCT-1995; 95MO-US12358.
XX      09-JAN-1995; 95US-0370190.
XX      30-SEP-1994; 94US-0316423.
XX      (LIFE) LIFE TECHNOLOGIES INC.
XX      Chatterjee DK, Hughes AJ;
XX      WPI; 1996-209362/21.
XX      P-PSDB; AAR98144.
XX      Thermotoga neopolitana DNA polymerase and thermostable mutants
XX      useful in DNA sequencing and amplification reactions
XX      Example 10: Fig 5A-B; 67pp; English.
XX      A DNA fragment (AAT30049) codes for the C-terminal portion (AAR98144
XX      of the 100 kDa thermostable DNA polymerase (AAR98143) of Thermotoga
XX      neopolitana (Tne), including the O helical region. The
XX      polymerase is a thermostable enzyme suitable for use in DNA
XX      sequencing and amplification reactions. It can be obtd. by
XX      obtd. from transformed host cells, e.g. E. coli, carrying the
XX      DNA polymerase gene. Mutants having improved enzymatic properties
XX      may be produced (see also AAR98144-47).
XX      Sequence 694 BP; 230 A; 135 C; 179 G; 150 T; 0 other;
XX      Alignment Score:
XX      Pred. No.: 0.495 Length: 694
XX      Score: 47.00 Matches: 10
XX      Percent Similarity: 90.91% Conservative: 0
XX      Best Local Similarity: 90.91% Mismatches: 1
XX      Query Match: 95.92% Indels: 0
XX      DB: 17 Gaps: 0
XX      US-09-823-649A-5 (1-11) x AAT30049 (1-694)
XX      QY 1 LeuSerValArgLeuGly**ProValIysCln 11
XX      Db 233 CTTTCGTGAGACTTGGAATACCGGTTAAAGAA 265
XX      RESULT 2
XX      AAT60031
XX      ID AAT60031 standard; DNA; 694 BP.
XX      AAT60031;
XX      09-JUN-1997 (first entry)
XX      Thermotoga neopolitana DNA polymerase C-terminal region DNA.
XX      DNA polymerase; Tne; exonuclease; DNA amplification; DNA sequencing;
XX      enzyme engineering; protein engineering; mutagenesis; ss.
XX      Thermotoga neopolitana strain DSM 5068.
XX      WO9709451-A1.
XX      13-MAR-1997.
XX      06-SEP-1996; 96MO-US14189.
XX      14-AUG-1996; 96US-0525057.
XX      08-SEP 1995; 95US-0525057.

```

```

PR      02-OCT-1995:   95US-05371397.
PR      02-OCT-1995:   95US-0537400.
PR      21-DEC-1995;    95US-0576759.

FA      (CHAV/) CHATTERJEE D. K.
PA      (HUGH/) HUGHES A. J.
XX      Chatterjee DK, Hughes AJ,
XX      WPI: 1997-192925/17.
DR      P-PsDB: AAM13862, AAM13863, AAM13864.
XX      XX
XX      Mutant DNA polymerase, esp. from Thermotoga - has reduced
PT      exonuclease activity or diexonucleotide discrimination, useful
PT      for DNA amplification, synthesis and sequencing
XX      XX
PS      Example 5; Fig 5; 120bp; English.
XX      XX
CC      A DNA sequence (AAT60031) encodes the C-terminal region, including
CC      the O-helix, of the Thermotoga neapolitana (Tne) DNA polymerase.
CC      It was obt'd. from a genomic DNA expression library by screening E.
CC      coli transformants for thermostable DNA polymerase activity. When
CC      the sequence is translated it does not produce the entire amino
CC      acid sequence of the polymerase due to frameshift errors generated
CC      during sequencing. However, an amino acid sequence was obt'd. by
CC      translating all 3 reading frames (AAM13862-64), comparing these with
CC      known sequences, and splicing the Tne sequence together to form the
CC      sequence given in AAM1363. A complete coding sequence for the
CC      polymerase is given in AAT60032.
XX      XX
SQ      Sequence 694 BP; 230 A; 136 C; 178 G; 150 T; 0 other;
Alignment Scores:
Pred. No.:          0.495           Length:         694
Score:              47.00           Matches:        10
Percent Similarity: 90.91%          Conservative:   0
Best Local Similarity: 90.91%       Mismatches:     1
Query Match:        95.92%          Indels:         0
DB:                 18             Gaps:           0
US-09-823-649A-5 (1-11) x AAT60031 (1-694)
QY      1 LeuservAlarglenGly**ProvallySGLu 11
Db      233 CTTCCTCGACACTGCAGATACCCGTAAAGAA 265
RESUIN 3
ID      AA288767
XX      AA288767 standard; cDNA; 694 BP.
XX      XX
AC      AA288767;
XX      XX
DT      18-MAY-2000 (first entry)
XX      XX
DE      T. neapolitana The polymerase C-terminal cDNA.
XX      XX
KW      DNA polymerase; PolI; O-helix; DNA sequencing; amplification; mutant;
KM      thermostability; Tne; ss.
XX      XX
OS      Thermotoga neapolitana.
XX      XX
Key      Location/Qualifiers
FT      CDS
FT      1..694
FT      /tag= "a"
FT      /product= "Tne"
FT      /note= "partial sequence, no start codon given"
XX      US6015668-A.
XX      PD      18-JAN-2000.
XX      EP      06-SEP-1996;    96US-0706706.
XX      XX

```

```

PR 02-OCT-1995; 95US-05373397.
PR 30-SEP-1994; 94US-0318423.
PR 09-JAN-1995; 95US-0370190.
PR 08-SEP-1995; 95US-0525057.
PR 02-OCT-1995; 95US-0537400.
PR 21-DEC-1995; 95US-0576759.
PR 14-AUG-1996; 96US-0689818.
XX
PA (LIFE-) LIFE TECHNOLOGIES INC.
PA Hughes AJ, Chatterjee DK;
XX
XX MPI: 2000-181138/16.
DR P-PSDB; AAY51572.
XX
PT DNA sequencing, amplification and synthesis methods, using mutant forms
PT of Thermotoga neapolitana polymerase with reduced exonuclease activity
PT or reduced discrimination against dinucleotides.
XX
PS Example 5; Column 51-54; 65pp; English.
XX
CC This invention describes novel methods for sequencing DNA, amplifying
CC double-stranded DNA or synthesizing DNA which use a mutant Thermotoga
CC neapolitana DNA polymerase (I) or its fragments. (I) are used for
CC sequencing, amplification and synthesis of DNA by essentially standard
CC methods, also for DNA labeling and for synthesis of cDNA from mRNA
CC templates. (I) have higher thermostability than known enzymes (e.g.
CC 50% retention of activity after 60 min at 90 plusoc), and may be free
CC of exonuclease activity and/or ability to discriminate against
CC dinucleotides (ddNTP). They permit sequencing at high temperature;
CC provide strong, clear sequences (increased signal-to-noise ratio); can
CC sequence more than 500 bases; reduce false stops in the sequence ladder;
CC reduce template requirements 10-fold (when using 35S labeling); and also
CC requirements for primer and labeled terminator. They increase sequencing
CC efficiency with structure-containing or GC-rich templates; reduce film
CC exposure times; do not generate sequencing artefacts with homopolymeric
CC sequences and allow quantitation of bases in heterologous mixtures. This
CC sequence encodes the Thermotoga neapolitana The polymerase carboxyl
CC terminal which is described in the method of the invention.
CC
XX
SQ Sequence 694 BP; 229 A; 136 C; 179 G; 150 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.495 Length: 694
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 21 Gaps: 0
XX
US-09-823-649A-5 (1-11) x AA88767 (1-694)
XX
CY 1 LeuSerValArgIeuGly***ProValIlyGlu 11
Db 233 CTTTCTGTGAGACTGTGATACCGGTAAAGAA 265
XX
RESULT 4
ID AAT58155 standard; DNA; 1485 BP.
XX
XX AAT58155;
XX
XX 15-APR-1997 (first entry)
XX
DE Thermotoga neapolitana DNA polymerase Tne M284 delB gene.
XX
XX DNA polymerase; Tne; thermostable enzyme; protein engineering;
XX thermostability; DNA sequencing; amplification; PCR;
XX polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
XX
XX Synthetic.
XX
XX WO9641014-A1.

```

PD	19-DEC-1996.	96WO-US09641.	
XX			
PF	07-JUN-1996;	96WO-US09641.	
XX			
PR	31-MAY-1996;	96US-0656664.	
PR	07-JUN-1995;	95US-0484661.	
XX			
PA	(PROM-) PROMEGA CORP.		
XX			
PI	Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;		
PI	Slater MR, Stortis DR;		
DR	WPI: 1997-052374/05.		
DR	P-PsDB; AAM11525.		
XX			
PT	Thermotoga DNA polymerase from Thermotoga neapolitana - useful in		
PT	DNA sequencing or for amplification of double stranded DNA		
XX			
PS	Example 4: Page 174-175; 200pp; English.		
XX			
CC	A DNA sequence (AAT58155) encodes deletion mutant Tne M284 delB		
CC	(AAM11515) of the DNA polymerase (see also AAM11513) of Thermotoga		
CC	neapolitana. The deletion mutant lacks the N-terminal 283		
CC	amino acids from the 5' exonuclease domain of the native enzyme		
CC	and also has amino acid residues 309-424 of the 3' exonuclease		
CC	domain deleted. The DNA construct was obtd. by restriction		
CC	digestion of pM284-derived DNA (see also AAT58135). When expressed		
CC	in E. coli, the M284 delB showed no detectable polymerase activity		
CC	In contrast, a Tne M284 mutant (see also AAM11514) was active.		
XX			
SQ	Sequence 1485 BP; 490 A; 312 C; 365 G; 318 T; 0 other;		
	Alignment Scores:		
	Pred. No.:	1.17	Length: 1485
	Score:	47.00	Matches: 10
	Percent Similarity:	90.91%	Conservative: 0
	Best Local Similarity:	90.91%	Mismatches: 1
	Query Match:	95.92%	Indels: 0
	DB:	18	Gaps: 0
	US-09-823-649A-5 (1-11) x AAT58155 (1-1485)		
OY	1 LeuSeVaIArLeuGly**ProValYsGln 11		
DB	1024 CTTCTCTGACATCTGCATATCCGCTTAAGCA 1056		
RESULT 5			
ID	AAT58145		
AC	AAT58145 standard; DNA: 1719 BP.		
XX			
XX	AAT58145;		
TF	15-APR-1997 (first entry)		
DE			
XX			
XX	Thermotoga neapolitana DNA polymerase Tne M23 gene.		
KW	DNA polymerase; Tne; thermotoga enzyme; protein engineering;		
KW	Thermotoga; Tne; DNA sequencing; amplification; PCR;		
KW	polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss		
OS	Synthetic.		
XX			
XX	MO9641014-A1.		
PD	19-DEC-1996.		
XX			
XX	07-JUN-1996;	96WO-US09641.	
XX			
PR	31-MAY-1996;	96US-0656664.	
PR	07-JUN-1995;	95US-0484661.	
XX			
XX	(PROM-) PROMEGA CORP.		

XX Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storis DR;
XX WPI: 1997-052374/05.
DR P-PSDB: AAM11524.
XX
PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
PI DNA sequencing or for amplification of double stranded DNA
XX
XX Example 4; Page 169-172; 200pp; English.
XX
CC A DNA sequence (AAT58145) encodes deletion mutant Tne M323 (AAM11524)
CC of the DNA polymerase (see also AAM11513) of Thermotoga neapolitana.
CC The M323 has an artificially created methionine at amino acid
CC position 323 as the initiator methionine in the 3' exonuclease
CC domain of the enzyme. When expressed in E. coli, the M323 showed
CC no detectable polymerase activity. In contrast, a Tne M284 mutant
CC (see also AAM11514) was active.
XX
SQ Sequence 1719 BP; 553 A; 367 C; 413 G; 386 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.38 Length: 1719
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservatve: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 18 Gaps: 0
XX
US-09-823-649a-5 (1-11) x AAT58145 (1-1719)
QY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 1255 CTTTCTGTGAGACTGTGAAATACCGGTTAAAGAA 1287
XX
RESULT 6
AAT58144
ID AAT58144 standard; DNA: 1737 BP.
XX
AC AAT58144;
XX
DT 15-APR-1997 (first entry)
XX
DE Thermotoga neapolitana DNA polymerase Tne M316 gene.
XX
KW DNA polymerase; Tne; thermostable enzyme; protein engineering;
KW thermostability; DNA sequencing; amplification; PCR;
KW polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
XX
OS Synthetic.
XX
PN WO9641014-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09641.
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storis DR;
XX
DR WPI: 1997-052374/05.
DR P-PSDB: AAM11523.
XX
PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
PI DNA sequencing or for amplification of double stranded DNA
XX
PS Example 3; Page 133-135; 200pp; English.

XX
CC A DNA sequence (AAT58144) encodes deletion mutant Tne M316 (AAM11523)
CC of the DNA polymerase (see also AAM11513) of Thermotoga neapolitana.
CC The M316 has an artificially created methionine at amino acid
CC position 316 as the initiator methionine. It was created by
CC introducing a Met residue at position 316 (and a corresponding
CC NcoI site) via site-directed mutagenesis of the native Tne gene
CC (see also AAT58134). When expressed in E. coli from a 17 or Lrc
CC promoter, the M316 produced no detectable polymerase activity.
CC A Tne M284 mutant (see also AAM11514) was active.
XX
SQ Sequence 1737 BP; 553 A; 374 C; 417 G; 393 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.4 Length: 1737
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservatve: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 18 Gaps: 0
XX
US-09-823-649a-5 (1-11) x AAT58144 (1-1737)
QY 1 LeuSerValArgLeuGly**ProValLysGlu 11
Db 1276 CTTTCTGTGAGACTGTGAAATACCGGTTAAAGAA 1308
XX
RESULT 7
AAT58135
ID AAT58135 standard; DNA: 1833 BP.
XX
AC AAT58135;
XX
DT 15-APR-1997 (first entry)
XX
DE Thermotoga neapolitana DNA polymerase Tne M284 gene.
XX
KW DNA polymerase; Tne; thermostable enzyme; protein engineering;
KW thermostability; DNA sequencing; amplification; PCR;
KW polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
XX
OS Synthetic.
XX
PN WO9641014-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09641.
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storis DR;
XX
DR WPI: 1997-052374/05.
DR P-PSDB: AAM11514.
XX
PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
PI DNA sequencing or for amplification of double stranded DNA
XX
PS Claim 20; Page 128-131; 200pp; English.
XX
CC A DNA sequence (AAT58135) in plasmid pM284 codes for modified DNA
CC polymerase Tne M284 (AAM11514) in which amino acids 1-283 of the
CC native Thermotoga neapolitana enzyme (see also AAM11513) are deleted
CC and the naturally occurring Met-284 residue becomes the initiating
CC methionine. It was obtd. by restriction digestion of the native
CC gene (AAT58134). The deletion results in elimination of 5'
CC exonuclease activity. Further modification of the DNA in pM284
CC (see also AAT58136-43) yield mutant enzymes (AAM11515-22) that also

CC have reduced or eliminated 3' exonuclease activity. Such enzymes
CC are useful in DNA sequencing reactions and nucleic acid
CC amplifications, esp. PCR.

XX Sequence 1833 BP; 591 A; 394 C; 440 G; 408 T; 0 other;

Alignment Scores:

Pred. No.:	1.49	Length:	1833
Score:	47.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	95.92%	Indels:	0
DB:	18	Gaps:	0

US-09-823-649A-5 (1-11) x AAT58135 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValLysGlu 11

DB 1372 CTTCTCTGAGACCTTGGAATACCGGTTAAAGAA 1404

RESULT 8

ID AAT58136 standard; DNA: 1833 BP.

AC AAT58136;

DT 15-APR-1997 (first entry)

DE Thermotoga neapolitana DNA polymerase Tne M284(D323E) gene.

XX DNA polymerase; Tne; thermostable enzyme; protein engineering;

KW thermostability; DNA sequencing; amplification; PCR;

KW polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.

OS Synthetic.

XX mutation

FT 120 /tag= a /note= "base 120 (A) is C in the native Tne gene"

PN W09641014-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96MO-US09641.

PR 31-MAY-1996; 96US-0656664.

PA (PROM-) PROMEGA CORP.

PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;

PI Slater MR, Storls DR;

DR WPI: 1997-052374/05.

DR P-PSDB; AAM11515.

PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in

PT DNA sequencing or for amplification of double stranded DNA

XX Claim 20; Page 138-140; 200pp; English.

CC A DNA sequence (AAT58136) codes for mutant thermostable DNA polymerase

CC The M284(D323E) (AAM11515) in which amino acids 1-283 of the native

CC Thermotoga neapolitana enzyme (see also AAM11513) are deleted and

CC amino acid residue 323 (Asp) of native Tne is substd. by Glu. It

CC was obtd. by site-directed mutagenesis of pm284 DNA (see also

CC AAT58135). The mutated gene can be incorporated into a vector and

CC used for prodn. of Tne M284(D323E) in transformed host cells. The

CC mutant enzyme has no significant 5' or 3' exonuclease activity and

CC is suitable for use in DNA sequencing and amplification reactions,

CC e.g. PCR.

XX Sequence 1833 BP; 592 A; 393 C; 440 G; 408 T; 0 other;

Alignment Scores:

Pred. No.:	1.49	Length:	1833
Score:	47.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	95.92%	Indels:	0
DB:	18	Gaps:	0

US-09-823-649A-5 (1-11) x AAT58136 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValLysGlu 11

DB 1372 CTTCTCTGAGACCTTGGAATACCGGTTAAAGAA 1404

RESULT 9

ID AAT58137 standard; DNA: 1833 BP.

AC AAT58137;

DT 15-APR-1997 (first entry)

DE Thermotoga neapolitana DNA polymerase Tne M284(E325D) gene.

XX DNA polymerase; Tne; thermostable enzyme; protein engineering;

KW thermostability; DNA sequencing; amplification; PCR;

KW polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.

OS Synthetic.

XX mutation

FT 126 /tag= a /note= "base 126 (C) is A in the native Tne gene"

PN W09641014-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96MO-US09641.

PR 31 MAY 1996; 96US 0656664.

PA (PROM-) PROMEGA CORP.

PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;

PI Slater MR, Storls DR;

DR WPI: 1997-052374/05.

DR P-PSDB; AAM11516.

PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in

PT DNA sequencing or for amplification of double stranded DNA

XX Claim 20; Page 142-144; 200pp; English.

CC A DNA sequence (AAT58137) codes for mutant thermostable DNA polymerase

CC The M284(E325D) (AAM11516) in which amino acids 1-283 of the native

CC Thermotoga neapolitana enzyme (see also AAM11513) are deleted and

CC amino acid residue 325 (Glu) of native Tne is substd. by Asp. It

CC was obtd. by site-directed mutagenesis of pm284 DNA (see also

CC AAT58135). The mutated gene can be incorporated into a vector and

CC used for prodn. of Tne M284(E325D) in transformed host cells. The

CC mutant enzyme has no significant 5' or 3' exonuclease activity and

CC is suitable for use in DNA sequencing and amplification reactions,

CC e.g. PCR.

XX Sequence 1833 BP; 590 A; 395 C; 440 G; 408 T; 0 other;

Alignment Scores:

Pred. No.:	1.49	Length:	1833
Score:	47.00	Matches:	10
Percent Similarity:	90.91%	Conservative:	0
Best Local Similarity:	90.91%	Mismatches:	1
Query Match:	95.92%	Indels:	0
DB:	18	Gaps:	0

US-09-823-649a-5 (1-11) x AAT58137 (1-1833)

QY 1 LeuSerValArgIeuGly***ProValIysGlu 11
 |||||
 DB 1372 CTTTCTGTGAGACTTGGAATACCGTTAAGAA 1404

RESULT 10
 AAT58138
 ID AAT58138 standard; DNA: 1833 BP.
 AC AAT58138;
 AT 15-APR-1997 (first entry)
 DE Thermotoga neapolitana DNA polymerase Tne M284(Y464F) gene.
 KW DNA polymerase; Tne; thermostable enzyme; protein engineering;
 KW thermostability; DNA sequencing; amplification; PCR;
 KW polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT mutation 542
 FT /*tag= "a"
 FT /note= "base 542 (T) is A in the native Tne gene"
 FT

W09641014-A1.
 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US09641.
 XX 31-MAY-1996; 96US-0656664.
 PR 07-JUN-1995; 95US-0484661.
 PA (PROM-) PROMEGA CORP.
 XX Holchakova E, Hartnelt JR, Huang F, Miller KM, Otto P;
 PI Slater MR, Storts DR;
 DR WPI: 1997-052374/05.
 DR P-PSDB: AAM1517.
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 PS Claim 20; Page 147-149; 200pp; English.
 XX A DNA sequence (AAT58138) codes for mutant thermostable DNA polymerase
 CC The M284(Y464F) (AAM1517) in which amino acids 1-283 of the native
 CC Thermotoga neapolitana enzyme (see also AAM1513) are deleted and
 CC amino acid residue 464 (Tyr) of native Tne is substid. by Phe. It
 CC was obtd. by site-directed mutagenesis of pm284 DNA (see also
 CC AAT58135). The mutated gene can be incorporated into a vector and
 CC used for prodn. of Tne M284(Y464F) in transformed host cells. The
 CC mutant enzyme has no significant 5' exonuclease activity and
 CC reduced 3' exonuclease activity. It is suitable for use in DNA
 CC sequencing and amplification reactions, e.g. PCR.
 XX

SO Sequence 1833 BP; 590 A; 393 C; 440 G; 410 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.49 Length: 1833
 Score: 47.00 Matches: 10
 Percent Similarity: 90.91%
 Best Local Similarity: 90.91%
 Query Match: 95.92%
 Indels: 0

Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 95.92% Indels: 0
 DB: 18 Gaps: 0

US-09-823-649a-5 (1-11) x AAT58138 (1-1833)

QY 1 LeuSerValArgIeuGly***ProValIysGlu 11
 |||||
 DB 1372 CTTTCTGTGAGACTTGGAATACCGTTAAGAA 1404

RESULT 11
 AAT58139
 ID AAT58139 standard; DNA: 1833 BP.
 AC AAT58139;
 AT 15-APR-1997 (first entry)
 DE Thermotoga neapolitana DNA polymerase Tne M284(D468N) gene.
 KW DNA polymerase; Tne; thermostable enzyme; protein engineering;
 KW thermostability; DNA sequencing; amplification; PCR;
 KW polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT mutation 553
 FT /*tag= "a"
 FT /note= "base 553 (A) is G in the native Tne gene"
 FT

W09641014-A1.
 19-DEC-1996.
 XX 07-JUN-1996; 96WO-US09641.
 XX 31-MAY-1996; 96US-0656664.
 PR 07-JUN-1995; 95US-0484661.
 PA (PROM-) PROMEGA CORP.
 XX Holchakova E, Hartnelt JR, Huang F, Miller KM, Otto P;
 PI Slater MR, Storts DR;
 DR WPI: 1997-052374/05.
 DR P-PSDB: AAM1518.
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 PS Claim 20; Page 151-154; 200pp; English.
 XX A DNA sequence (AAT58139) codes for mutant thermostable DNA polymerase
 CC The M284(Y464F) (AAM1518) in which amino acids 1-283 of the native
 CC Thermotoga neapolitana enzyme (see also AAM1513) are deleted and
 CC amino acid residue 468 (Asp) of native Tne is substid. by Asn. It
 CC was obtd. by site-directed mutagenesis of pm284 DNA (see also
 CC AAT58135). The mutated gene can be incorporated into a vector and
 CC used for prodn. of Tne M284(D468N) in transformed host cells. The
 CC mutant enzyme has no significant 5' exonuclease activity and
 CC reduced 3' exonuclease activity. It is suitable for use in DNA
 CC sequencing and amplification reactions, e.g. PCR.
 XX

SO Sequence 1833 BP; 591 A; 394 C; 439 G; 409 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.49 Length: 1833
 Score: 47.00 Matches: 10
 Percent Similarity: 90.91%
 Best Local Similarity: 90.91%
 Query Match: 95.92%
 Indels: 0

DB: 18 Gaps: 0

US-09-823-649A-5 (1-11) x AAT58139 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValIysGlu 11
 |||||
 DB 1372 CTTTCTGTGAGACTGTGAAATACCGGTTAAAGAA 1404

RESULT 12

AAT58140
 ID AAT58140 standard; DNA; 1833 BP.
 AC AAT58140;
 XX
 DT 15-APR-1997 (first entry)
 DE Thermotoga neapolitana DNA polymerase Tne M284(D323A) gene.
 XX
 KM DNA polymerase; Tne; thermostable enzyme; protein engineering;
 KM thermostability; DNA sequencing; amplification; PCR;
 KM polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
 XX
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT mutation 119
 FT /tag= a
 FT /note= "base 119 (C) is A in the native Tne gene"

XX PN W09641014-A1.
 XX PD 19-DEC-1996.
 XX PE 07-JUN-1996; 96MO-US09641.
 XX PR 31-MAY-1996; 96US-0656664.
 XX PR 07-JUN-1995; 95US-0484661.
 XX
 PA (PROM-) PROMEGA CORP.
 PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
 PI Slater MR, Storts DR.
 XX
 DR WPI: 1997-052374/05.
 DR P-PSDB: AAW11519.
 XX
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 XX
 PS Claim 20: Page 156-158; 200pp; English.

CC A DNA sequence (AAT58140) codes for mutant thermostable DNA polymerase
 CC The M284(D323E) (AAW11519) in which amino acids 1-283 of the native
 CC Thermotoga neapolitana enzyme (see also AAW11513) are deleted and
 CC amino acid residue 323 (Asp) of native Tne is substd. by Ala. It
 CC was obtd. by site-directed mutagenesis of pm284 DNA (see also
 CC AAT58135). The mutated gene can be incorporated into a vector and
 CC used for prodn. of Tne M284(D323A) in transformed host cells. The
 CC mutant enzyme is suitable for use in DNA sequencing and
 CC amplification reactions, e.g. PCR.
 XX

SQ Sequence 1833 BP; 590 A; 395 C; 440 G; 408 T; 0 other;

Alignment Scores:
 Pred. No.: 1.49 Length: 1833
 Score: 47.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 95.92% Indels: 0
 DB: 18 Caps: 0

US-09-823-649A-5 (1-11) x AAT58140 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValIysGlu 11
 |||||
 DB 1372 CTTTCTGTGAGACTGTGAAATACCGGTTAAAGAA 1404

RESULT 13

AAT58141
 ID AAT58141 standard; DNA; 1833 BP.
 AC AAT58141;
 XX
 DT 15-APR-1997 (first entry)
 DE Thermotoga neapolitana DNA polymerase Tne M284(D389A) gene.
 XX
 KM DNA polymerase; Tne; thermostable enzyme; protein engineering;
 KM thermostability; DNA sequencing; amplification; PCR;
 KM polymerase chain reaction; exonuclease; Thermotoga neapolitana; ss.
 XX
 OS Synthetic.
 XX
 FT Key Location/Qualifiers
 FT mutation 317..318
 FT /tag= a
 FT /note= "bases 317-318 (CG) are AC in the native Tne gene"

XX PN W09641014-A1.
 XX PD 19-DEC-1996.
 XX PE 07-JUN-1996; 96MO-US09641.
 XX PR 31-MAY-1996; 96US-0656664.
 XX PR 07-JUN-1995; 95US-0484661.
 XX
 PA (PROM-) PROMEGA CORP.
 PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
 PI Slater MR, Storts DR.
 XX
 DR WPI: 1997-052374/05.
 DR P-PSDB: AAW11520.
 XX
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 XX
 PS Claim 20: Page 161-163; 200pp; English.

CC A DNA sequence (AAT58141) codes for mutant thermostable DNA polymerase
 CC The M284(D323E) (AAW11520) in which amino acids 1-283 of the native
 CC Thermotoga neapolitana enzyme (see also AAW11513) are deleted and
 CC amino acid residue 389 (Asp) of native Tne is substd. by Ala. It
 CC was obtd. by site-directed mutagenesis of pm284 DNA (see also
 CC AAT58135). The mutated gene can be incorporated into a vector and
 CC used for prodn. of Tne M284(D389A) in transformed host cells. The
 CC mutant enzyme is suitable for use in DNA sequencing and
 CC amplification reactions, e.g. PCR.
 XX

SQ Sequence 1833 BP; 590 A; 394 C; 441 G; 408 T; 0 other;

Alignment Scores:
 Pred. No.: 1.49 Length: 1833
 Score: 47.00 Matches: 10
 Percent Similarity: 90.91% Conservative: 0
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 95.92% Indels: 0
 DB: 18 Caps: 0

US-09-823-649A-5 (1-11) x AAT58141 (1-1833)

QY 1 LeuSerValArgLeuGly***ProValIysGlu 11
 |||||
 DB 1372 CTTTCTGTGAGACTGTGAAATACCGGTTAAAGAA 1404

Tue Jan 21 14:46:07 2003

us-09-823-649a-5.rng

Page 7

```

RESULT 14
AAT58142
ID AAT58142 standard; DNA: 1833 BP.
XX
AC AAT58142;
XX
DT 15-APR-1997 (first entry)
XX
DE Thermotoga neopolitana DNA polymerase Tne M284(D323A,D389A) gene.
XX
KM DNA polymerase; Tne; thermostable enzyme; protein engineering;
KM thermostability; DNA sequencing; amplification; PCR;
KM polymerase chain reaction; exonuclease; Thermotoga neopolitana; ss.
OS Synthetic.
XX
FH Key Location/Qualifiers
FH mutation 119
FH /tag= a
FH /note= "base 119 (C) is A in the native Tne gene"
FH mutation 317..318
FH /tag= b
FH /note= "bases 317-318 (CC) are AC in the native Tne
FH gene"
XX
PN M09641014-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09641.
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storts DR;
XX
DR WPI: 1997-052374/05.
DR P-PSDB: AAM11521.
XX
PT Thermostable DNA polymerase from Thermotoga neopolitana - useful in
PT DNA sequencing or for amplification of double stranded DNA
XX
PS Claim 20; Page 165-167; 200pp; English.
XX
CC A DNA sequence (AAT58142) codes for mutant thermostable DNA polymerase
CC Tne M284(D323E,D389A) (AAM11521) in which amino acids 1-283 of the
CC native Thermotoga neopolitana enzyme (see also AAM11513) are deleted
CC and amino acid residue 323 (ASP) of native Tne is substd. by Ala,
CC and Asp-389 by Ala. It was obtd. by site-directed mutagenesis of
CC PM284 DNA (see also AAT58135). The mutated gene can be incorporated
CC into a vector and used for prodn. of the triple mutant in
CC transformed host cells. The triple mutant has no significant 5' or
CC 3' exonuclease activity and a half-life at 97.5 deg C of 22 min
CC compared with 5 min for native Tne and 8 min for native Tag. It
CC is suitable for use in DNA sequencing and amplification reactions,
CC e.g. PCR.
XX
SQ Sequence 1833 BP; 589 A; 395 C; 441 G; 408 T; 0 other;

Alignment Scores:
Pred. No.: 1.49 Length: 1833
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservatve: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 18 Gaps: 0

US-09-823-649a-5 (1-11) x AAT58142 (1-1833)

```

```

OY 1 LeuSerValArgLeuGly***ProValLysGlu 11
DB 1372 CTTTCCTCAGCACTTGGAATACCGCTTAAGAA 1404
RESUP 15
AAT58143
ID AAT58143 standard; DNA: 1833 BP.
XX
AC AAT58143;
XX
DT 15-APR-1997 (first entry)
XX
DE Tne Quad polymerase gene.
XX
KM DNA polymerase; Tne; Quad polymerase; thermostable enzyme;
KM protein engineering; thermostability; DNA sequencing; PCR;
KM polymerase chain reaction; amplification; exonuclease;
KM Thermotoga neopolitana; ss.
OS Synthetic.
XX
FH Key Location/Qualifiers
FH mutation 119
FH /tag= a
FH /note= "base 119 (C) is A in the native Tne gene"
FH mutation 317..318
FH /tag= b
FH /note= "bases 317-318 (CC) are AC in the native Tne
FH gene"
XX
PN M09641014-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09641.
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storts DR;
XX
DR WPI: 1997-052474/05.
DR P-PSDB: AAM11522.
XX
PT Thermostable DNA polymerase from Thermotoga neopolitana - useful in
PT DNA sequencing or for amplification of double stranded DNA
XX
PS Claim 20; Page 180-182; 200pp; English.
XX
CC A DNA sequence (AAT58143) codes for mutant thermostable DNA polymerase
CC Tne Quad (AAM11522) in which amino acids 1-283 of the native
CC Thermotoga neopolitana enzyme (see also AAM11513) are deleted and
CC amino acid residue 3 amino acid substns. (D323A, D389A, F730Y) are
CC made in the 3' exonuclease domain of the enzyme. It was obtd. by
CC site-directed mutagenesis of DNA (AAT58142) coding for a triple
CC mutant. The mutated gene can be incorporated into a vector and
CC used for prodn. of Tne Quad in transformed host cells for use
CC in DNA amplification reactions and esp. for radioactive sequencing
XX
SQ Sequence 1833 BP; 590 A; 395 C; 440 G; 408 T; 0 other;

Alignment Scores:
Pred. No.: 1.49 Length: 1833
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservatve: 0
Best Local Similarity: 90.91% Mismatches: 1

```

Query Match: 95.92% Indels: 0
DB: 18 Caps: 0

US-09-823-649A-5 (1-11) x AAT58143 (1-1833)

OY 1 LeuSeValArgLeuGly**ProValIysGlu 11
|||||
Db 1372 CTTTCTGTGAGACTGTGATACCGGTAAAGAA 1404

Search completed: January 15, 2003, 12:55:39
Job time : 137.143 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame.plus.p2n model

Run on: January 15, 2003, 12:36:21 ; Search time 26 Seconds
(without alignments)
129.748 Million cell updates/sec

Title: US-09-823-649a-5
Perfect score: 49
Sequence: 1 LSVRLGCPYKE 11

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15333881 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODE=frame-p2n model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US09823649.rnmat.14012003.151002.29127/app.query.fasta_1.1393

-DB=Issued Patents_NA -OPMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -ICORCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=diosun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.@CGN.1.1.35 @rnmat.14012003.151002.29127 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MAP -LARGESQUPRY -NRG SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUTS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	95.9	694	2	US-08-537-400-15
2	47	95.9	694	2	US-08-706-702-17
3	47	95.9	694	3	US-08-706-706-17
4	47	95.9	1485	3	US-08-484-661A-38
5	47	95.9	1485	3	US-08-656-664-38
6	47	95.9	1485	5	PCT-US96-09641-38
7	47	95.9	1716	3	US-08-484-661A-36
8	47	95.9	1716	3	US-08-656-664-36
9	47	95.9	1716	3	PCT-US96-09641-36
10	47	95.9	1737	3	US-08-484-661A-10
11	47	95.9	1737	3	US-08-656-664-10
12	47	95.9	1737	5	PCT-US96-09641-10

13	47	95.9	1833	3	US-08-484-661A-7	Sequence 7, Appl 1
14	47	95.9	1833	3	US-08-484-661A-15	Sequence 15, Appl 1
15	47	95.9	1833	3	US-08-484-661A-18	Sequence 18, Appl 1
16	47	95.9	1833	3	US-08-484-661A-22	Sequence 22, Appl 1
17	47	95.9	1833	3	US-08-484-661A-25	Sequence 25, Appl 1
18	47	95.9	1833	3	US-08-484-661A-28	Sequence 28, Appl 1
19	47	95.9	1833	3	US-08-484-661A-32	Sequence 32, Appl 1
20	47	95.9	1833	3	US-08-484-661A-34	Sequence 34, Appl 1
21	47	95.9	1833	3	US-08-656-664-7	Sequence 7, Appl 1
22	47	95.9	1833	3	US-08-656-664-15	Sequence 15, Appl 1
23	47	95.9	1833	3	US-08-656-664-18	Sequence 18, Appl 1
24	47	95.9	1833	3	US-08-656-664-22	Sequence 22, Appl 1
25	47	95.9	1833	3	US-08-656-664-25	Sequence 25, Appl 1
26	47	95.9	1833	3	US-08-656-664-28	Sequence 28, Appl 1
27	47	95.9	1833	3	US-08-656-664-32	Sequence 32, Appl 1
28	47	95.9	1833	3	US-08-656-664-34	Sequence 34, Appl 1
29	47	95.9	1833	3	US-08-656-664-53	Sequence 53, Appl 1
30	47	95.9	1833	5	PCT-US96-09641-7	Sequence 7, Appl 1
31	47	95.9	1833	5	PCT-US96-09641-15	Sequence 15, Appl 1
32	47	95.9	1833	5	PCT-US96-09641-18	Sequence 18, Appl 1
33	47	95.9	1833	5	PCT-US96-09641-22	Sequence 22, Appl 1
34	47	95.9	1833	5	PCT-US96-09641-25	Sequence 25, Appl 1
35	47	95.9	1833	5	PCT-US96-09641-28	Sequence 28, Appl 1
36	47	95.9	1833	5	PCT-US96-09641-32	Sequence 32, Appl 1
37	47	95.9	1833	5	PCT-US96-09641-34	Sequence 34, Appl 1
38	47	95.9	1833	5	PCT-US96-09641-53	Sequence 53, Appl 1
39	47	95.9	2682	1	US-07-971-819A-1	Sequence 1, Appl 1
40	47	95.9	2682	1	US-07-971-434-3	Sequence 3, Appl 1
41	47	95.9	2682	1	US-08-475-231-1	Sequence 1, Appl 1
42	47	95.9	2682	1	US-08-458-819-3	Sequence 3, Appl 1
43	47	95.9	2682	2	US-08-706-702-2	Sequence 2, Appl 1
44	47	95.9	2682	3	US-08-484-661A-1	Sequence 1, Appl 1
45	47	95.9	2682	3	US-08-706-706-2	Sequence 2, Appl 1

ALIGNMENTS

RESULT 1
US-08-537-400-15
Sequence 15, Application US/08537400
Patent No. 5939301
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Cloned DNA Polymerases From Thermotoga
TITLE OF INVENTION: neapolitana And Mutants Thereof
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: STEPHEN KESSLER, COMUSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/537,400
FILING DATE: 02-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,423
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/370,190
FILING DATE: 09-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2800002
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 694 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: 2..694
US-08-537-400-15

Alignment Scores:
Pred. No.: 0.0822 Length: 694
Score: 47.00 Matches: 10
Percent Similarity: 90.918 Conservative: 0
Best Local Similarity: 90.918 Mismatches: 1
Query Match: 95.928 Indels: 0
DB: 2 Gaps: 0

US-09-823-6494-5 (1-11) x US-08-537-400-15 (1-694)

QY 1 LeuSerValArgLeuGly**ProValIysGlu 11
|||||
Db 233 CTTCTGTGAGACTTGGATACCGGTAAAGAA 265

RESULT 2

US-08-706-702-17
Sequence 17, Application US/08706702
Patent No. 5948614
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA polymerases from Thermotoga
NUMBER OF INVENTION: Matrima and Mutants Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,702
CLASSIFICATION: 435
FILING DATE: 06-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942,2800006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600

TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 694 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..691
US-08-706-702-17

Alignment Scores:
Pred. No.: 0.0822 Length: 694
Score: 47.00 Matches: 10
Percent Similarity: 90.918 Conservative: 0
Best Local Similarity: 90.918 Mismatches: 1
Query Match: 95.928 Indels: 0
DB: 2 Gaps: 0

US-09-823-6494-5 (1-11) x US-08-706-702-17 (1-694)

QY 1 LeuSerValArgLeuGly**ProValIysGlu 11
|||||
Db 233 CTTCTGTGAGACTTGGATACCGGTAAAGAA 265

RESULT 3

US-08-706-706-17
Sequence 17, Application US/08706706
Patent No. 6015668
GENERAL INFORMATION:
APPLICANT: Hughes, A. John
TITLE OF INVENTION: Cloned DNA polymerases from Thermotoga and
NUMBER OF INVENTION: Matrima and Mutants Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,706
CLASSIFICATION: 435
FILING DATE: 06-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/370,190
FILING DATE: 01-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,423
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:
NAME: Esmoud, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2800005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 694 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..691
US-08-706-706-17

Alignment Scores:
Pred. No.: 0.0822 Length: 694
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: Caps: 0

US-09-823-649a-5 (1-11) x US-08-706-706-17 (1-694)

OY 1 LeuserValArgLeucly***ProValIysGlu 11
Db 233 CTTCTCTGACACTTGGAATACCGTTAAAGAA 265

RESULT 4
US-08-484-661A-38
Sequence 38, Application US/08484661A
Patent No. 6001645
GENERAL INFORMATION:
APPLICANT: SLATER, MICHAEL R.
APPLICANT: HARTNETT, JAMES R.
APPLICANT: HUANG, FEN
APPLICANT: BOICHAKOVA, ELENA
TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,661A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-01175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1485 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
US-08-484-661A-38

Alignment Scores:
Pred. No.: 0.2 Length: 1485
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: Caps: 0

US-09-823-649a-5 (1-11) x US-08-484-661A-38 (1-1485)

OY 1 LeuserValArgLeucly***ProValIysGlu 11
Db 1024 CTTCTCTGACACTTGGAATACCGTTAAAGAA 1056

RESULT 5
US-08-656-664-38
Sequence 39, Application US/08656664
Patent No. 6077664
GENERAL INFORMATION:
APPLICANT: Slater, Michael R.
APPLICANT: Huang, Fen
APPLICANT: Hartnett, James R.
APPLICANT: Boichakova, Elena
APPLICANT: Storts, Douglas R.
APPLICANT: Otto, Paul
TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,664
FILING DATE: 31-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-02185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1485 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1482
US-08-656-664-38

Alignment Scores:
Pred. No.: 0.2 Length: 1485
Score: 47.00 Matches: 10

Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-5 (1-11) x US-08-656-664-38 (1-1485)

Qy 1 LeuSerValArgLeuGly***ProValLysGlu 11
Db 1024 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1056

RESULT 6

PCT-US96-09641-38
; Sequence 38, Application PC/TUS9609641

GENERAL INFORMATION:

APPLICANT: Slater, Michael R.

APPLICANT: Huang, Fen

APPLICANT: Hartnett, James R.

APPLICANT: Bolchakova, Elena

APPLICANT: Storts, Douglas R.

APPLICANT: Otto, Paul

TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM

TITLE OF INVENTION: THERMOTOGA NEAPOLITANA

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States Of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/09641

FILING DATE: 31-MAY-1996

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: PRMG-02185

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 1485 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1482

PCT-US96-09641-38

Alignment Scores:

Pred. No.: 0.2 Length: 1485
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 5 Gaps: 0

US-09-823-649A-5 (1-11) x PCT-US96-09641-38 (1-1485)

Qy 1 LeuSerValArgLeuGly***ProValLysGlu 11
Db 1024 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1056

RESULT 7
US-08-484-661A-36
; Sequence 36, Application US/08484661A
; Patent No. 6001645

GENERAL INFORMATION:

APPLICANT: SLATER, MICHAEL R.

APPLICANT: HARTNETT, JAMES R.

APPLICANT: HUANG, FEN

APPLICANT: BOLCHAKOVA, ELENA

TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES

TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: MEDLEN & CARROLL

STREET: 220 MONTGOMERY STREET, SUITE 2200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,661A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: PRMG-01175

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 397-8338

TELEFAX: (415) 705-8410

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 1716 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1713

US-08-484-661A-36

Alignment Scores:

Pred. No.: 0.237 Length: 1716
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-5 (1-11) x US-08-484-661A-36 (1-1716)

Qy 1 LeuSerValArgLeuGly***ProValLysGlu 11
Db 1255 CTTTCTGTGAGACTTGGAATACCGGTTAAAGAA 1287

RESULT 8

US-08-656-664-36
; Sequence 36, Application US/0865664
; Patent No. 6077664

GENERAL INFORMATION:

APPLICANT: Slater, Michael R.

APPLICANT: Huang, Fen

APPLICANT: Hartnett, James R.

APPLICANT: Bolchakova, Elena

APPLICANT: Storts, Douglas R.

APPLICANT: Otto, Paul

TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM

;; TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
;; NUMBER OF SEQUENCES: 57
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medien & Carroll
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States Of America
;; ZIP: 94104
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/656,664
;; FILING DATE: 31-MAY-1996
;;
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ingolia, Diane E.
;; REGISTRATION NUMBER: 40,027
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;;
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1716 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1713
;;
;; US-08-656-664-36
;;
Alignment Scores:
Pred. No.: 0.237 Length: 1716
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conserva: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0
US-09-823-649a-5 (1-11) x US-08-656-664-36 (1-1716)
QY 1 LeuserValArgIeugly**ProValIysCiu 11
DB 1255 CTTCTGTGAGACTTGATACCGGTAAAGAA 1287
;;
RESULT 9
PCT-US96-09641-36
;; Sequence 36 Application PC/TUS9609641
;; GENERAL INFORMATION:
;; APPLICANT: Slater, Michael R.
;; APPLICANT: Huang, Fen
;; APPLICANT: Hartnett, James R.
;; APPLICANT: Bolchakova, Elena
;; APPLICANT: Storts, Douglas R.
;; APPLICANT: Otto, Paul
;; TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
;; TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
;; NUMBER OF SEQUENCES: 57
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medien & Carroll
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States Of America
;; ZIP: 94104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US96/09641
;; FILING DATE: 31-MAY-1996
;;
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ingolia, Diane E.
;; REGISTRATION NUMBER: 40,027
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;;
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1716 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1713
;;
PCT-US96-09641-36
;;
Alignment Scores:
Pred. No.: 0.237 Length: 1716
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conserva: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 5 Gaps: 0
US-09-823-649a-5 (1-11) x PCT-US96-09641-36 (1-1716)
QY 1 LeuserValArgIeugly**ProValIysCiu 11
DB 1255 CTTCTGTGAGACTTGATACCGGTAAAGAA 1287
;;
RESULT 10
US-08-484-661A-10
;; Sequence 10 Application US/08484661A
;; Patent No. 6001645
;; GENERAL INFORMATION:
;; APPLICANT: Slater, Michael R.
;; APPLICANT: Huang, Fen
;; APPLICANT: Hartnett, James R.
;; APPLICANT: Bolchakova, Elena
;; APPLICANT: Storts, Douglas R.
;; APPLICANT: Otto, Paul
;; TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
;; TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
;; NUMBER OF SEQUENCES: 51
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medien & Carroll
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States Of America
;; ZIP: 94104
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/484,661A
;; FILING DATE: 07-JUN-1995
;;
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ingolia, Diane E.
;; REGISTRATION NUMBER: 40,027
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;;
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1716 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..1713
;;
PCT-US96-09641-36

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1734
US-08-484-661A-10

Alignment Scores:
Pred. No.: 0.24 Length: 1737
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-5 (1-11) x US-08-484-661A-10 (1-1737)
QY 1 LeuserValArgLeugly***ProValTysGlu 11
Db 1276 CTTTCGTGACGCTTGGAATACCGTTAAAGAA 1308

RESULT 11
US-08-656-664-10
Sequence 10, Application US/08656664
Patent No. 6077664
GENERAL INFORMATION:
APPLICANT: Slater, Michael R.
APPLICANT: Huang, Fen
APPLICANT: Hartnett, James R.
APPLICANT: Bolchakova, Elena
APPLICANT: Storts, Douglas R.
APPLICANT: Otto, Paul
TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,664
FILING DATE: 31-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-02185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS
LOCATION: 1..1734
US-08-656-664-10

Alignment Scores:
Pred. No.: 0.24 Length: 1737
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-5 (1-11) x US-08-656-664-10 (1-1737)
QY 1 LeuserValArgLeugly***ProValTysGlu 11
Db 1276 CTTTCGTGACGCTTGGAATACCGTTAAAGAA 1308

RESULT 12
PCT-US96-09641-10
Sequence 10, Application PC/TUS9609641
GENERAL INFORMATION:
APPLICANT: Slater, Michael R.
APPLICANT: Huang, Fen
APPLICANT: Hartnett, James R.
APPLICANT: Bolchakova, Elena
APPLICANT: Storts, Douglas R.
APPLICANT: Otto, Paul
TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSER: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09641
FILING DATE: 31-MAY-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-02185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1737 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1734
PCT-US96-09641-10

Alignment Scores:
Pred. No.: 0.24 Length: 1737
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 5 Gaps: 0

US-09-823-649A-5 (1-11) x PCR-US96-09641-10 (1-1737)

QY 1 LeuserValArgLeugly***ProValLysgu 11
|||||
Db 1276 CTTCTGTGACACTTGCAATACCGTTAAAGAA 1408

RESULT 13

US-08-484-661A-7
; Sequence 7, Application US/08484661A
; Patent No. 6001645
; GENERAL INFORMATION:
; APPLICANT: SLATER, MICHAEL R.
; APPLICANT: HARTNETT, JAMES R.
; APPLICANT: HUANG, FEN
; APPLICANT: BOLCHAKOVA, ELENA
; TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
; TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,661A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: PRMG-01175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1833 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1830
US-08-484-661A-7
Alignment Scores:
Pred. No.: 0.256 Length: 1833
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0
US-09-823-649A-5 (1-11) x US-08-484-661A-7 (1-1833)
QY 1 LeuserValArgLeugly***ProValLysgu 11
|||||
Db 1372 CTTCTGTGACACTTGCAATACCGTTAAAGAA 1404
RESULT 14
US-08-484-661A-15
; Sequence 15, Application US/08484661A
; Patent No. 6001645
; GENERAL INFORMATION:

APPLICANT: SLATER, MICHAEL R.
APPLICANT: HARTNETT, JAMES R.
APPLICANT: HUANG, FEN
APPLICANT: BOLCHAKOVA, ELENA
TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,661A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-01175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1830
US-08-484-661A-15

Alignment Scores:
Pred. No.: 0.256 Length: 1833
Score: 47.00 Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best Local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-5 (1-11) x US-08-484-661A-15 (1-1833)

QY 1 LeuserValArgLeugly***ProValLysgu 11
|||||
Db 1372 CTTCTGTGACACTTGCAATACCGTTAAAGAA 1404

RESULT 15

US-08-484-661A-18
; Sequence 18, Application US/08484661A
; Patent No. 6001645
; GENERAL INFORMATION:
; APPLICANT: SLATER, MICHAEL R.
; APPLICANT: HARTNETT, JAMES R.
; APPLICANT: HUANG, FEN
; APPLICANT: BOLCHAKOVA, ELENA
; TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
; TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,661A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INCOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-01175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ. ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1830
US-08-484-661A-18

Alignment Scores:
Pred. No.: 0.256 length: 1833
Score: 47.00* Matches: 10
Percent Similarity: 90.91% Conservative: 0
Best local Similarity: 90.91% Mismatches: 1
Query Match: 95.92% Indels: 0
DB: 3 Gaps: 0

US-09-823-649A-5 (1-11) x US-08-484-661A-18 (1-1833)

QY 1 LeuSerValArgLeuGly**ProValIlyGlu 11
|||||
Db 1372 CTTCTGTGAGACTTGGAATACCGGTAAAGAA 1404

Search completed: January 15, 2003, 12:58:49
Job time : 29 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:15:24 : Search time 12.1429 Seconds

(without alignments)
87.086 Million cell updates/sec

Title: US-09-823-649a-5

Perfect score: 49

Sequence: 1 LSVRLGKPKVE 11

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	95.9	893 2 E72232	DNA-directed DNA p
2	38	77.6	324 2 D81785	conserved hypotet
3	38	77.6	324 2 G81207	sugar isomerase, K
4	37	75.5	441 2 A13607	NADH2 dehydrogenas
5	37	75.5	451 2 E72398	hypothetical prote
6	36	73.5	477 1 ANRR	angiotensin precur
7	36	73.5	753 1 D72660	probable aldehyde
8	35	71.4	315 2 F70203	xylose operon regu
9	34	69.4	277 2 E75361	conserved hypotet
10	34	69.4	334 2 F64925	ying protein - Bsc
11	34	69.4	334 2 A99927	hypothetical prote
12	34	69.4	334 2 B85775	hypothetical prote
13	34	69.4	356 2 D55020	recombinational prot
14	34	69.4	379 2 C88445	protein C26E.6 li
15	34	69.4	557 2 S61154	hypothetical prote
16	34	69.4	676 2 D95249	transcription regu
17	34	69.4	676 2 A98114	conserved hypotet
18	34	69.4	997 2 I54390	phosphoinositide-s
19	34	69.4	1096 2 SE2358	inositol 1,4,5-tri
20	34	69.4	1208 2 T00362	hypothetical prote
21	34	69.4	1251 2 C82721	conserved hypotet
22	33	67.3	380 2 C75016	na+/H+ antiporter
23	33	67.3	621 2 S73155	hypothetical prote
24	33	67.3	1490 2 JC5145	DNA (cytosine-5-)-
25	33	65.3	183 2 B90092	hypothetical prote
26	32	65.3	311 2 A93603	oxidoreductase (EC
27	32	65.3	317 2 S33450	apolipoprotein E -
28	32	65.3	341 1 VMBP80	major capsid prote
29	32	65.3	363 2 A75294	A/G-specific aden

30	32	65.3	429 2 C97500	hypothetical prote
31	32	65.3	429 2 AG2718	MFS permease (limp
32	32	65.3	473 2 A84943	membrane protein [
33	32	65.3	684 2 G90473	hypothetical prote
34	32	65.3	688 2 B90275	hypothetical prote
35	32	65.3	772 2 AH0597	ferrous iron trans
36	32	65.3	773 1 A36932	iron(II) transport
37	32	65.3	773 2 C91160	ferrous iron trans
38	32	65.3	773 2 B86006	ferrous iron trans
39	32	65.3	964 1 T04325	probable ATP-depen
40	32	65.3	1034 2 S76134	hypothetical prote
41	32	65.3	1165 2 T38867	probable DNA-direc
42	32	65.3	1226 1 S65593	adenosine deaminas
43	32	65.3	1284 2 T50993	hypothetical prote
44	31	63.3	107 2 AD2489	hypothetical prote
45	31	63.3	107 2 AG2490	hypothetical prote

ALIGNMENTS

RESULT 1

E72232 DNA-directed DNA polymerase I - Thermotoga maritima (strain MSB)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: E72232

R.Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: E72232

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-893 <ARN>

A:Cross-references: GB:AE001805; GB:AE00512; NID:g4982180; PTDN:AAD36686.1; PTD:g498

A:Experimental source: strain MSB

C:Genetics:

A:Gene: TM1619

C:Superfamily: DNA-directed DNA polymerase I

Query Match 95.9%; Score 47; DB 2; Length 893;

Best Local Similarity 90.9%; Pred. No. 0.17;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGKPKVE 11

Db 741 LSVRLGKPKVE 751

RESULT 2

D81785

conserved hypothetical protein NMA2135 [imported] - Neisseria meningitidis (strain Z2

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: D81785

R.Parthill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: D81785

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-824 <PK>

A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PTDN:CA85347.1; PTD:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2135

C:Superfamily: probable ATP-binding protein gutQ; CBS homology

Query Match 77.6%; Score 38; DB 2; Length 324;
Best Local Similarity 70.0%; Pred. No. 4.4;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SVRLGXPYKE 11
:|||||:
DB 213 AVRLGTPLEKE 222

RESULT 3

G81207
sugar isomerase, Kpsf/GutQ family NMB0352 [Imported] - Neisseria meningitidis (strain MC
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: G81207
R:Reichel, H.; Saunders, N.J.; Heibelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M.
Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappelli, R.; Ve
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A/Reference number: A81000; MIMD:2017555; PMID:10710307
A/Accession: G81207
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-324 <TEF>
A/Cross-references: GB:AE002392; GB:AE002098; NID:97225573; PIDN:AAE40795.1; PID:9722557
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NMB0352
C/Superfamily: probable ATP-binding protein gutQ; CBS homology

Query Match 72.6%; Score 38; DB 2; Length 324;
Best Local Similarity 70.0%; Pred. No. 4.4;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SVRLGXPYKE 11
:|||||:
DB 213 AVRLGTPLEKE 222

RESULT 4

A13607
NADH2 dehydrogenase (EC 1.6.99.3) [Imported] - Brucella melitensis (strain 16M)
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Jun-2002
C/Accession: A13607
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A/Reference number: AD3252; PMID:11756688
A/Accession: A13607
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-441 <KOR>
A/Cross-references: GB:AE008918; PIDN:AAU54028.1; PID:g17984981; GSPDB:GN00191
A/Experimental source: strain 16M
C/Genetics:
A/Gene: BME110786
A/Map position: 11
C/Superfamily: NADH dehydrogenase
C/Keywords: oxidoreductase

Query Match 75.5%; Score 37; DB 2; Length 441;
Best Local Similarity 63.6%; Pred. No. 9.7;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSVRLGXPYKE 11
:|||||:
DB 247 VEVRLGIPYKD 257

RESULT 5
E72398
hypothetical protein TM0244 - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C/Accession: E72398
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A/Reference number: A72200; MIMD:99287316; PMID:10360571
A/Accession: E72398
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-451 <ARN>
A/Cross-references: GB:AE001708; GB:AE000512; NID:94980740; PIDN:AAU35335.1; PID:9498
A/Experimental source: strain MSB8
C/Genetics:
A/Gene: TM0244

Query Match 75.5%; Score 37; DB 2; Length 451;
Best Local Similarity 63.6%; Pred. No. 9.9;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSVRLGXPYKE 11
:|||||:
DB 313 LVRLGTTPYKD 323

RESULT 6

ANRT
angiotensin precursor - rat
N/Contents: angiotensin I; angiotensin II; angiotensin III
C/Species: Rattus norvegicus (Norway rat)
C/Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 18-Jun-1999
C/Accession: A93945; A90456; A01251
R:Ohkubo, H.; Kageyama, R.; Ujihara, M.; Hirose, T.; Inayama, S.; Nakamishi, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 2196-2200, 1983
A/Title: Cloning and sequence analysis of cDNA for rat angiotensinogen.
A/Reference number: A93945; MIMD:83169849; PMID:6572971
A/Accession: A93945
A/Molecule type: mRNA
A/Residues: 1-477 <OHK>

A/Cross-references: GB:J00094; GB:J00704; NID:9702912; PIDN:AAE98779.1; PID:9702914
R:Bohnik, J.; Clausen, E.; Strosberg, D.; Treney, J.P.; Menard, J.; Corvol, P.
Biochemistry 20, 7010-7015, 1981
A/Title: Rat angiotensinogen and Des(angiotensin)angiotensinogen: purification, char
A/Reference number: A90456; MIMD:82091819; PMID:6797467
A/Accession: A90456
A/Molecule type: protein
A/Residues: 25-41 <BOH>

C/Comment: Angiotensin I is released from angiotensinogen by renin, which is secreted
e I (angiotensin-converting enzyme), primarily in the lungs.
C/Comment: The release of the amino-terminal residue (Asp-25) from angiotensin I and
sp-1]angiotensin I is converted to angiotensin III by dipeptidyl carboxypeptidase I.
C/Comment: Angiotensinogen is synthesized in the liver and secreted into the plasma.
ung.

C/Superfamily: antithrombin III
C/Keywords: blood pressure control; glycoprotein; liver; plasma; vasoconstrictor
F:1-24/DNA: signal sequence #status predicted <SIG>
F:25-477/Product: angiotensinogen #status predicted <SP>
F:25-34/Product: angiotensin I #status experimental <PPI>
F:25-33/Product: angiotensin II #status experimental <PP2>
F:26-33/Product: angiotensin III #status experimental <PP3>
F:295,319/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.5%; Score 36; DB 1; Length 477;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSVRLGXPYKE 11
:|||||:

Db 148 LQVILGVPVK 158

RESULT 7

D72660

probable aldehyde oxidoreductase APE0708 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 16-Jun-2000

C:Accession: D72660
K:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takai-
awa, H.; Takamaya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: D72660

A:Molecule type: DNA

A:Residues: 1-753 <KAW>

A:Cross-references: DDBJ:AP000060; NID:95104188; PIDN:BAW9664.1; PID:95104369

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0708

C:Superfamily: carbon-monoxide dehydrogenase large chain

Query Match

73.5%; Score 36; DB 1; Length 753;

Best Local Similarity 70.0%; Pred. No. 27;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVK 10

1:11111111

Db 252 LAMRLGXPVK 261

RESULT 8

F70203

xylose operon regulatory protein (xylR-2) homolog - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: F70203

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vglt,
; Bowman, C.; Garland, S.; Fujili, C.; Colton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: F70203

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-315 <KLE>

A:Cross-references: GB:AE001181; GB:AE000783; NID:92688772; PIDN:AAC67180.1; PID:9268877

A:Experimental source: strain B31

C:Superfamily: glucose kinase

F:61-188/Domain: glucose kinase homology <GKH>

Query Match 71.4%; Score 35; DB 2; Length 315;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVK 11

1:11111111

Db 190 VAMRLGXPVK 200

RESULT 9

E75361

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: E75361

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Ulteback, T.; Zalewski, C.; Mc

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75361

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-277 <WH1>

A:Cross-references: GB:AE002014; GB:AE000513; NID:96459494; PIDN:AAF11284.1; PID:9645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1728

A:Map position: 1

C:Superfamily: ydaA protein

Query Match

69.4%; Score 34; DB 2; Length 277;

Best Local Similarity 63.6%; Pred. No. 25;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSVRLGXPVK 11

1:11111111

Db 242 LPARLGPVKD 252

RESULT 10

F64925

yng protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: F64925

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: F64925

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-334 <BLAT>

A:Cross-references: GB:AE000263; GB:U00096; NID:q1787966; PIDN:AAC74748.1; PID:q17879

A:Experimental source: strain K-12, substrain MG155

C:Genetics:

A:Gene: yng

C:Superfamily: conserved hypothetical protein b0819

Query Match

69.4%; Score 34; DB 2; Length 334;

Best Local Similarity 77.8%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SVRLGXPVK 10

11111111

Db 223 SVRLGXPVK 231

RESULT 11

A99927

hypothetical protein KCS2385 (imported) - Escherichia coli (strain 0157:H7, substrain

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: A99927

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C

gasawara, N.; Tasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and g

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: A99927

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <HAV>

A:Cross-references: GB:BA000007; PIDN:BA835808.1; PID:q13361852; GSDB:GNO0154

A:Experimental source: strain 0157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: ECS2385

C:Superfamily: conserved hypothetical protein b0819

Query Match 69.4%; Score 34; DB 2; Length 334;
 Best Local Similarity 77.8%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SVRLGXPVK 10
 ||| |||
 Db 223 SVRLGXPVK 231

RESULT 12

E85775
 hypothetical protein ynhg [imported] - Escherichia coli (strain O157:H7, substrain ED193)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: E85775
 R:Perna, M.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grothbeck, C.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potlamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: E85775
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <STO>
 A:Cross-references: GB:AE005174; NID:g12515680; PIDD:AA65665.1; GSPDB:GN00145; UMGF:227
 A:Experimental source: strain O157:H7, substrain ED1933
 C:Genetics:
 A:Gene: ynhg
 C:Superfamily: conserved hypothetical protein b0819

Query Match 69.4%; Score 34; DB 2; Length 334;
 Best Local Similarity 77.8%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 SVRLGXPVK 10
 ||| |||
 Db 223 SVRLGXPVK 231

RESULT 13

D55020
 recombination protein recA - Thermotoga maritima (strain MS8)
 N:Alternate names: DNA repair protein; recombinase A
 C:Species: Thermotoga maritima
 C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 02-Feb-2001
 C:Accession: D55020; D72201
 R:Wentur, J.G.; Wong, D.M.; Ortiz, B.; Tong, J.; Reichert, F.; Geland, D.H.
 J. Biol. Chem. 269, 25928-25935, 1994
 A:Title: Cloning, sequencing, and expression of RecA proteins from three distantly relat
 A:Reference number: A55020; MUID:95014407; PMID:7929298
 A:Accession: D55020
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-356 <NEF>
 A:Cross-references: GB:L2425; NID:g385169; PIDD:AAA27417.1; PID:g385170
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Heft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: D72201
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-356 <ARN>
 A:Cross-references: GB:AE001823; GB:AE000512; NID:g4982441; PIDD:AAJ36921.1; PID:g498244
 A:Experimental source: strain MS8
 C:Genetics:
 A:Gene: TM1859
 C:Superfamily: recombination protein recA
 C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
 F142-147/Region: nucleotide-binding motif B

F:74/Binding site: ATP (lys) #status predicted
 Query Match 69.4%; Score 34; DB 2; Length 356;
 Best Local Similarity 54.5%; Pred. No. 33;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LSVRLGXPVK 11
 : ||| |||
 Db 225 MEVRLGXPVK 235

RESULTS 14

C88445
 protein C26E6.6 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: C88445
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A:Accession: C88445
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-379 <STO>
 A:Cross-references: GB:chr_III; PIDD:AAA21160.1; PID:g532807; GSPDB:GN00021; CESP:C26
 A:Note: similar to ribosomal protein L3
 C:Genetics:
 A:Gene: C26E6.6
 A:Map position: 3

Query Match 69.4%; Score 34; DB 2; Length 379;
 Best Local Similarity 77.8%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VRLGXPVK 11
 ||| |||
 Db 162 VRLGXPVK 170

RESULTS 15

S61154
 hypothetical protein YDR358w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein D9476.2
 C:Species: Saccharomyces cerevisiae
 C:Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
 C:Accession: S61154
 R:Du, X.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of S. cerevisiae cosmid 9476.
 A:Reference number: S61148
 A:Accession: S61154
 A:Molecule type: DNA
 A:Residues: 1-557 <UOZ>
 A:Cross-references: EMBL:U28372; NID:g849170; PID:g849172; GSPDB:GN00004; MIPS:YDR358
 A:Gene: SGD:GAL1; MIPS:YDR358w
 A:Cross-references: SGD:S0002766
 A:Map position: 48
 C:Superfamily: Schizosaccharomyces pombe hypothetical protein SWH25H2.1oc

Query Match 69.4%; Score 34; DB 2; Length 557;
 Best Local Similarity 60.0%; Pred. No. 51;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LSVRLGXPVK 10
 : ||| |||
 Db 523 ISVRLGXPVK 532

Search completed: January 15, 2003, 12:37:23

Tue Jan 21 14:46:09 2003

us-09-823-649a-5.rpr

Page 5

Job time : 16.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 09:02:09 ; Search time 28.1429 Seconds

(without alignments)
52.083 Million cell updates/sec

Title: US-09-823-649A-5
Perfect score: 49
Sequence: 1 LSVRLGXPKVE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	95.9	11	AA847795	Native DNA polymer
2	47	95.9	11	AA848266	Native DNA polymer
3	47	95.9	11	AA848267	Native DNA polymer
4	47	95.9	230	AA898144	The DNA polymerase
5	47	95.9	230	AA898145	The DNA polymerase
6	47	95.9	230	AA813863	The DNA polymerase
7	47	95.9	230	AA813863	The DNA polymerase
8	47	95.9	494	AA81525	Thermotoga neopol
9	47	95.9	572	AA81524	Thermotoga neopol
10	47	95.9	578	AA81523	Thermotoga neopol

11	47	95.9	609	19	AA83971	The DNA polymerase
12	47	95.9	610	18	AA81514	Thermotoga neopol
13	47	95.9	610	18	AA81515	Thermotoga neopol
14	47	95.9	610	18	AA81516	Thermotoga neopol
15	47	95.9	610	18	AA81517	Thermotoga neopol
16	47	95.9	610	18	AA81518	Thermotoga neopol
17	47	95.9	610	18	AA81519	Thermotoga neopol
18	47	95.9	610	18	AA81520	Thermotoga neopol
19	47	95.9	610	18	AA81521	Thermotoga neopol
20	47	95.9	610	18	AA81522	Thermotoga neopol
21	47	95.9	610	19	AA83970	The Quad polymerase
22	47	95.9	610	19	AA83971	The DNA polymerase
23	47	95.9	610	19	AA83972	The DNA polymerase
24	47	95.9	611	13	AA83973	The DNA polymerase
25	47	95.9	611	13	AA83974	The DNA polymerase
26	47	95.9	674	19	AA83969	The DNA polymerase
27	47	95.9	674	19	AA84000	The DNA polymerase
28	47	95.9	674	19	AA84001	The DNA polymerase
29	47	95.9	674	19	AA84002	The DNA polymerase
30	47	95.9	677	19	AA84003	The DNA polymerase
31	47	95.9	700	19	AA83972	The DNA polymerase
32	47	95.9	755	13	AA83350	Mutant thermotabi
33	47	95.9	774	20	AA83376	Mutant thermotabi
34	47	95.9	821	13	AA83349	Mutant thermotabi
35	47	95.9	857	13	AA83347	Mutant thermotabi
36	47	95.9	875	13	AA83348	Mutant thermotabi
37	47	95.9	893	13	AA821830	Sequence of a ther
38	47	95.9	893	13	AA823146	Mutant thermotabi
39	47	95.9	893	18	AA813899	Thermotoga neopol
40	47	95.9	893	18	AA813900	Thermotoga neopol
41	47	95.9	893	18	AA813901	Thermotoga neopol
42	47	95.9	893	18	AA813902	Thermotoga neopol
43	47	95.9	893	18	AA81513	Thermotoga neopol
44	47	95.9	893	19	AA878758	The DNA polymerase
45	47	95.9	893	19	AA83974	The DNA polymerase

ALIGNMENTS

RESULT 1	
ID	AA847795
AA847795	standard; peptide; 11 AA.
AC	AA847795;
XX	
XX	
DT	25-MAR-2002 (first entry)
XX	
DE	Native DNA polymerase motif #5.
XX	
XX	
KW	DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KW	transverse transcription reaction; fluorescein; cyanine; thermoactive;
KW	dye; amplification.
XX	
OS	Thermotoga sp.
XX	
FH	Key
FT	Misc-difference
FT	7
XX	Location/Qualifiers
XX	/label= Val, Ile
PN	
XX	EP1152062-A2.
PD	07-NOV-2001.
XX	
XX	
XX	12-APR-2001; 2001EP-0109341.
PF	
XX	
XX	18-APR-2000; 2000US-198336P.
PR	
XX	
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
PA	
XX	
XX	Smith ES, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;
PI	Schoenbrunner NJ, Wang AM;
XX	

DR WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse

PJ transcriptase polymerase chain reaction amplification of a mixture

PS using a mutant thermoactive DNA polymerase -

XX

XX Claim 2; Page 4; 23pp; English.

XX

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the

CC native forms of motifs derived from DNA polymerases used in the method

CC of the invention. The method for reverse transcribing an RNA, comprises

CC treating a transverse transcription reaction mixture which comprises the

CC RNA, a primer, a divalent cation, and a mutant thermoactive DNA

CC polymerase. These motifs represent a conserved region which affects the

CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with

CC fluorescein and cyanine family dyes. The method of the invention is

CC useful in reverse transcription/amplification reactions. The method

CC provides improved reverse transcription efficiency at lower enzyme

CC concentrations relative to previous high temperature reverse

CC transcription methods, and provide a wider range of usable salt

CC concentrations. The use of the mutant DNA polymerase provides faster

CC reverse transcription extension rates, and consequently less time is

CC needed for the reaction. This motif is derived from DNA polymerases from

CC Thermotoga species maritima and neopolitana.

XX

SQ Sequence 11 AA:

Query Match 95.9%; Score 47; DB 23; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0033;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 LSVRLGXPKYE 11

|||||

DB 1 LSVRLGXPKYE 11

RESULT 2

AAM48266

ID AAM48266 standard; Peptide; 11 AA.

XX

AC AAM48266;

XX

DE 25-MAR-2002 (first entry)

XX

DE Native DNA polymerase motif #15.

XX

KW DNA polymerase; reverse transcription; primer; divalent cation; mutant;

KW transverse transcription reaction; fluorescein; cyanine; thermoactive;

KW dye; amplification.

XX

OS Thermotoga maritima.

XX

PN EP1152062-A2.

XX

PD 07-NOV-2001.

XX

PF 12-APR-2001; 2001EP-0109341.

XX

PR 18-APR-2000; 2000US-198336P.

XX

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX

PI Smith BS, Elfstrom CM, Gelfand DH, Higuchi RG, Myers TW;

PI Schoenbumer NJ, Wang AM;

XX

WPI; 2002-076891/11.

PT Reverse transcribing an RNA, comprises performing a reverse

PJ transcriptase polymerase chain reaction amplification of a mixture

PS using a mutant thermoactive DNA polymerase -

XX

XX Disclosure: Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of mollis derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermostable DNA
CC polymerase. These mollis represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method
CC provides improved reverse transcription efficiency at lower enzyme
CC concentrations relative to previous high temperature reverse
CC transcription methods, and provide a wider range of usable salt
CC concentrations. The use of the mutant DNA polymerase provides faster
CC reverse transcription extension rates, and consequently less time is
CC needed for the reaction.

CC XX

SO Sequence 11 AA;

Query Match 95.9%; Score 47; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0033;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Caps

Oy 1 LSVRLGVPVKE 11
Db 1 LSVRLGVPVKE 11

RESULT 3

AAM48267

ID AAM48267 standard; Peptide: 11 AA.

XX AAM48267;

XX 25-MAR-2002 (first entry)

DF Native DNA polymerase mollis #16.

DE

XX DNA polymerase; reverse transcription; primer; divalent cation; mutant;
KM Transverse transcription reaction; fluorescein; cyanine; thermostable;
RW dye; amplification.

XX Thermoloba neapolitana.

OS

XX EP1152062-A2.

PN 07-NCV-2001.

PD 12-APR-2001; 2001EP 0109341.

XX 18-APR-2000; 2000MS 198436P.

PR (HOF) HOFPMANN JA ROCHE & CO AG F.

PA Smith ES, Ellstrom CM, Gelland DU, Higuchi RG, Myers TW;
PI Schoenbrunner NJ, Wang AM;
PP WPI; 2002-076891/11.

XX

XX Reverse transcribing an RNA, comprises performing a reverse
PT transcriptase polymerase chain reaction amplification of a mixture
PP using a mutant thermostable DNA polymerase -

XX

PS Disclosure; Page 7; 23pp; English.

CC The sequences given in AAB47791-97 and AAM48259-AAM48270 represent the
CC native forms of mollis derived from DNA polymerases used in the method
CC of the invention. The method for reverse transcribing an RNA, comprises
CC treating a transverse transcription reaction mixture which comprises the
CC RNA, a primer, a divalent cation, and a mutant thermostable DNA
CC polymerase. These mollis represent a conserved region which affects the
CC DNA polymerase's ability to incorporate dideoxynucleotides labelled with
CC fluorescein and cyanine family dyes. The method of the invention is
CC useful in reverse transcription/amplification reactions. The method

CC provides improved reverse transcription efficiency at lower enzyme
 CC concentrations relative to previous high temperature reverse
 CC transcription methods, and provide a wider range of usable salt
 CC concentrations. The use of the mutant DNA polymerase provides faster
 CC reverse transcription extension rates, and consequently less time is
 CC needed for the reaction.

XX Sequence 11 AA;

Query Match 95.9%; Score 47; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0033; Mismatches 1; Indels 0; Gaps 0;

XX 1 LSVRLGIPVKE 11
 XX ||||| ||||
 Db 1 LSVRLGIPVKE 11

RESULT 4
 AAR98144
 ID AAR98144 standard; Protein; 230 AA.

XX AAR98144;

XX 15-AUG-1996 (first entry)

XX The DNA polymerase C-terminal region.

DE DNA polymerase; Tne; DNA sequencing; amplification;

KW thermostable enzyme.

XX Thermotoga neopolitana strain DSM 5068.

XX Key Location/Qualifiers

FT Region 63..72 /label= O-helical_region

XX WO9610640-A1.

XX 11-APR-1996.

XX 02-OCT-1995; 95WO-US12358.

XX 09-JAN-1995; 95US-0370190.

XX 30-SEP-1994; 94US-0316423.

XX (LIFE) LIFE TECHNOLOGIES INC.

XX Chatterjee DK, Hughes AJ;

XX WPI; 1996-209362/21.

XX N-PSDB: AAT30049.

XX Thermotoga neopolitana DNA polymerase and thermostable mutants

XX useful in DNA sequencing and amplification reactions

XX Example 10; Fig 5A-B; 67pp; English.

XX The C-terminal portion (AAR98144) of the 100 kDa thermostable DNA
 CC polymerase (AAR98143) of Thermotoga neopolitana (Tne) includes an
 CC O helical region. Tne polymerase is a thermostable enzyme suitable
 CC for use in DNA sequencing and amplification reactions. It can be
 CC obtd. from host cells, e.g. E. coli, carrying the DNA polymerase
 CC gene. Mutants having improved enzymatic properties may also be
 CC produced (see also AAR98144-47).

XX Sequence 230 AA;

Query Match 95.9%; Score 47; DB 17; Length 230;

Best Local Similarity 90.9%; Pred. No. 0.094; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGIPVKE 11

Db ||||| ||||
 78 LSVRLGIPVKE 88

RESULT 5
 AAR98145

ID AAR98145 standard; Protein; 230 AA.

XX AAR98145;

XX 15-AUG-1996 (first entry)

XX The DNA polymerase C-terminal region (R67Y mutant).

KW DNA polymerase; Tne; DNA sequencing; amplification;

XX thermostable enzyme; protein engineering; mutagenesis.

XX Synthetic.

XX Key Location/Qualifiers

FT Region 63..72 /label= O-helical_region

XX WO9610640-A1.

XX 11-APR-1996.

XX 02-OCT-1995; 95WO-US12358.

XX 09-JAN-1995; 95US-0370190.

XX 30-SEP-1994; 94US-0316423.

XX (LIFE) LIFE TECHNOLOGIES INC.

XX Chatterjee DK, Hughes AJ;

XX WPI; 1996-209362/21.

XX Thermotoga neopolitana DNA polymerase and thermostable mutants

XX useful in DNA sequencing and amplification reactions

XX Claim 50; Fig5A-B; 67pp; English.

XX Site-directed mutagenesis was used to substitute a tyrosine

XX residue for phenylalanine in the O-helical region of Thermotoga

XX neopolitana (Tne) DNA polymerase. This residue corresponds

XX to position 67 of a C-terminal fragment (see also AAR98144) of

XX the enzyme. As a result of the mutation, the polymerase becomes

XX non-discriminating against non-natural nucleotides such as

XX dideoxynucleotides. The mutant enzyme, which may also incorporate

XX a D323A mutation in the N-terminal portion (see AAR98146) or has

XX the 5'-3' exonuclease domain deleted (see AAR98147), is active and

XX heat stable, and can be used in DNA sequencing and amplification
 CC reactions.

XX Sequence 230 AA;

Query Match 95.9%; Score 47; DB 17; Length 230;

Best Local Similarity 90.9%; Pred. No. 0.094; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGIPVKE 11
 XX ||||| ||||
 Db 78 LSVRLGIPVKE 88

RESULT 6

AAW13863

ID AAW13863 standard; Protein; 230 AA.

XX AAW13863;

DT 09-JUN-1997 (first entry)

```

XX The DNA polymerase C-terminal region.
XX
XX DNA polymerase; Tne; exonuclease; DNA amplification; DNA sequencing;
XX enzyme engineering; protein engineering; mutagenesis.
XX
XX Thermotoga neapolitana strain DSM 5068.
XX
XX WO9709451-A1.
XX
XX 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-0514189.
XX
XX 14-AUG-1995; 96US-0525057.
XX 08-SEP-1995; 95US-0525057.
XX 02-OCT-1995; 95US-0537397.
XX 02-OCT-1995; 95US-0537400.
XX 21-DEC-1995; 95US-0576759.
XX
XX (CHAT/) CHATTERJEE D K.
XX (HUGH/) HUGHES A J.
XX
XX Chatterjee DK, Hughes AJ;
XX
XX WPI: 1997-192925/17.
XX
XX N-PSDB: AAT60031.
XX
XX
XX Mutant DNA polymerase, esp. from Thermotoga - has reduced
XX exonuclease activity or dideoxynucleotide discrimination, useful
XX for DNA amplification, synthesis and sequencing
XX
XX Example 5; Fig 5; 120pp; English.
XX
XX Polypeptide sequences (AAW13862-64) were deduced for all 3 reading
XX frames of a DNA sequence (AAT60031) encoding the C-terminal region,
XX including the O-helix, of the Thermotoga neapolitana (Tne) DNA
XX polymerase. Frameshift errors generated during sequencing of the
XX DNA prevented the entire amino acid sequence being deduced from a
XX single reading frame. The 3 sequences were compared with known
XX polymerase amino acid sequences and spliced to generate the Tne
XX sequence given in AAW13863. A complete amino acid sequence for Tne
XX polymerase is given in AAW13899.
XX
XX
XX Sequence 230 AA:
SQ
Query Match 95.9%; Score 47; DB 18; Length 230;
Best Local Similarity 90.9%; Pred. No. 0.094;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LSVRLGXPKVE 11
DB 78 LSVRLGXPKVE 88
RESULT 7
AAW51572
ID AAW51572 standard; Protein: 230 AA.
XX
XX AAW51572;
AC
XX 18-MAY-2000 (first entry)
XX
XX T. neapolitana Tne polymerase C-terminal protein fragment.
XX
XX DNA polymerase; Pol; O-helix; DNA sequencing; amplification; mutant;
XX thermostability; Tne.
XX
XX Thermotoga neapolitana.
XX
XX US6015668-A.
XX
XX 18-JAN-2000.
PD

```

```

XX
XX 06-SEP-1996; 96US-0706706.
XX
XX 02-OCT-1995; 95US-0537397.
XX 30-SEP-1994; 94US-0316423.
XX 09-JAN-1995; 95US-0370190.
XX 08-SEP-1995; 95US-0525057.
XX 02-OCT-1995; 95US-0537400.
XX 21-DEC-1995; 95US-0576759.
XX 14-AUG-1996; 96US-0689818.
XX
XX (LIFE-) LIFE TECHNOLOGIES INC.
XX
XX Hughes AJ, Chatterjee DK;
XX
XX WPI: 2000-181138/16.
XX
XX N-PSDB: AAZ88767.
XX
XX
XX DNA sequencing, amplification and synthesis methods, using mutant forms
XX of Thermotoga neapolitana polymerase with reduced exonuclease activity
XX or reduced discrimination against dideoxynucleotides -
XX
XX Example 5; Column 53-54; 65pp; English.
XX
XX This invention describes novel methods for sequencing DNA, amplifying
XX double-stranded DNA or synthesizing DNA which use a mutant Thermotoga
XX neapolitana DNA polymerase (I) or its fragments. (I) are used for
XX sequencing, amplification and synthesis of DNA by essentially standard
XX methods, also for DNA labeling and for synthesis of cDNA from mRNA
XX templates. (I) have higher thermostability than known enzymes (e.g.
XX 50% retention of activity after 60 min at 90 plusC), and may be free
XX of exonuclease activity and/or ability to discriminate against
XX dideoxynucleotides (ddNTP). They permit sequencing at high temperature;
XX provide strong, clear sequences (increased signal-to-noise ratio); can
XX sequence more than 500 bases; reduce false stops in the sequence ladder;
XX reduce template requirements 10-fold (when using 35S labeling); and also
XX requirements for primer and labeled terminator. They increase sequencing
XX efficiency with structure-containing or GC-rich templates; reduce film
XX exposure times; do not generate sequencing artefacts with homopolymeric
XX sequences and allow quantitation of bases in heterologous mixtures. This
XX sequence represents the Thermotoga neapolitana Tne polymerase carboxyl
XX terminal which is described in the method of the invention.
XX
XX
XX Sequence 230 AA:
SQ
Query Match 95.9%; Score 47; DB 21; Length 230;
Best Local Similarity 90.9%; Pred. No. 0.094;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LSVRLGXPKVE 11
DB 78 LSVRLGXPKVE 88
RESULT 8
AAW11525
ID AAW11525 standard; Protein: 494 AA.
XX
XX AAW11525;
AC
XX 15-APR-1997 (first entry)
XX
XX Thermotoga neapolitana DNA polymerase Tne K284 delB.
XX
XX DNA polymerase; Tne; thermostable enzyme; protein engineering;
XX thermostability; DNA sequencing; amplification; PCR;
XX polymerase chain reaction; exonuclease; Thermotoga neapolitana.
XX
XX Synthetic.
XX
XX WO9641014-A1.
XX
XX 19-DEC-1996.
PD

```

```

XX 07-JUN-1996; 96WO-US09641.
PF
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
XX (PROM-) PROMEGA CORP.
PA
XX Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storts DR;
XX
XX WPI: 1997-052374/05.
DR N-PSDB; AAT58145.
XX
PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
PT DNA sequencing or for amplification of double stranded DNA
XX
PS Example 4; Page 174-175; 200pp; English.
XX
CC Deletion mutant Tne M284 delB (AAW11525) of the Tne DNA polymerase
CC (see also AAW11513) of Thermotoga neapolitana lacks the N-terminal
CC 283 amino acids from the 5' exonuclease domain of the native enzyme
CC and also has amino acid residues 309-424 of the 3' exonuclease
CC domain deleted. It can be obtd. by expression of a DNA construct
CC (AAT58155) created by restriction digestion of pm284-derived DNA (see
CC also AAT58135). When expressed in E. coli, Tne M284 delB showed no
CC detectable polymerase activity. In contrast, a Tne M284 mutant
CC (see also AAW11514) was active.
XX
SQ Sequence 494 AA;
XX
Query Match 95.9%; Score 47; DB 18; Length 494;
Best Local Similarity 90.9%; Pred. No. 0.22;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGXPKYE 11
DB 342 LSVRLGIPYKE 352
XX
RESULT 9
AAW11524
ID AAW11524 standard; Protein: 572 AA.
XX
AC AAW11524;
XX
DT 15-APR-1997 (first entry)
XX
DE Thermotoga neapolitana DNA polymerase Tne M323 mutant.
XX
KW DNA polymerase; Tne; thermostable enzyme; protein engineering;
KW thermostability; DNA sequencing; amplification; PCR;
TW polymerase chain reaction; exonuclease; Thermotoga neapolitana.
XX
OS Synthetic.
OS
XX WO9641014-A1.
XX
XX 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09641.
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
XX (PROM-) PROMEGA CORP.
PA
XX Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storts DR;
XX
XX WPI: 1997-052374/05.
DR N-PSDB; AAT58145.
XX

```

```

PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
PT DNA sequencing or for amplification of double stranded DNA
XX
PS Example 4; Page 169-175; 200pp; English.
XX
CC Deletion mutant Tne M323 (AAW11524) of Thermotoga neapolitana DNA
CC polymerase Tne (see also AAW11513) has an artificially created
CC methionine at amino acid position 323 as the initiator methionine
CC in the 3' exonuclease domain of the enzyme. It can be expressed
CC using a DNA construct (AAT58145) created by restriction enzyme
CC digestion. When expressed in E. coli, Tne M323 produced no
CC detectable polymerase activity. In contrast, a Tne M284 mutant
CC (see also AAW11514) was active.
XX
SQ Sequence 572 AA;
XX
Query Match 95.9%; Score 47; DB 18; Length 572;
Best Local Similarity 90.9%; Pred. No. 0.26;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGXPKYE 11
DB 419 LSVRLGIPYKE 429
XX
RESULT 10
AAW11523
ID AAW11523 standard; Protein: 578 AA.
XX
AC AAW11523;
XX
DT 15-APR-1997 (first entry)
XX
DE Thermotoga neapolitana DNA polymerase Tne M316 mutant.
XX
KW DNA polymerase; Tne; thermostable enzyme; protein engineering;
KW thermostability; DNA sequencing; amplification; PCR;
KW polymerase chain reaction; exonuclease; Thermotoga neapolitana.
XX
OS Synthetic.
OS
XX WO9641014-A1.
XX
XX 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US09641.
XX
PR 31-MAY-1996; 96US-0656664.
PR 07-JUN-1995; 95US-0484661.
XX
XX (PROM-) PROMEGA CORP.
PA
XX Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
PI Slater MR, Storts DR;
XX
XX WPI: 1997-052374/05.
DR N-PSDB; AAT58144.
XX
XX Thermostable DNA polymerase from Thermotoga neapolitana - useful in
PT DNA sequencing or for amplification of double stranded DNA
XX
PS Example 3; Page 133-135; 200pp; English.
XX
CC Deletion mutant Tne M316 (AAW11523) of Thermotoga neapolitana DNA
CC polymerase Tne (see also AAW11513) has an artificially created
CC methionine at amino acid position 316 as the initiator methionine.
CC It can be expressed using a DNA construct (AAT58144) created by
CC introducing a Met residue at position 316 (and a corresponding
CC NcoI site) into the native gene (AAT58134) via site-directed
CC mutagenesis. When expressed in E. coli from a T7 or trc promoter,
CC Tne M316 produced no detectable polymerase activity. A Tne M284
CC mutant (see also AAW11514) was active.
XX

```

SQ Sequence 578 AA;
 Query Match 95.9%; Score 47; DB 18; Length 578;
 Best Local Similarity 90.9%; Pred. No. 0.26;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVE 11
 ||||| ||||
 Db 426 LSVRLGIPYKE 436

RESULT 11
 AAW83971
 ID AAW83971 standard; Protein; 609 AA.
 XX
 AC AAW83971;
 XX
 DT 08-FEB-1999 (first entry)
 XX
 DE The DNA polymerase mutant N'del284, D323A.
 XX
 KW The: O-helix; DNA polymerase; polymorphic; amplicon; thermostable;
 KM mutant; forensic; paternity testing; nucleic acid amplification; cancer;
 KM pathogen; genetic disease; cystic fibrosis; haemophilia; transplanting;
 KM Alzheimer's disease; screening; organ; diagnosis; plant breeding.
 XX
 OS Thermotoga neapolitana.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Msc-difference 39 /label= D323A
 FT /note="Wild-type Asp is replaced with Ala"
 XX
 PN WO9835060-A1.
 XX
 PD 13-AUG-1998.
 XX
 PF 09-FEB-1998; 98WO-US02791.
 XX
 PR 06-JAN-1998; 98US-0037393.
 PR 07-FEB-1997; 97US-0037393.
 XX
 PA (LIFE-) LIFE TECHNOLOGIES INC.
 XX
 PI Chatterjee DK, Solus J, Yang S;
 XX
 DR WPI: 1998-447254/38.
 XX
 PT New mutant polymerases that do not add non-templated 3'-nucleotides
 PT - specifically for analysis of hypervariable polymorphisms e.g. for
 PT forensic and paternity testing, disease diagnosis
 XX
 PS Claim 52: Page -, 187pp; English.
 XX
 CC Sequences AAW83969 to AAW84011 represent the (Thermotoga neapolitana)
 CC DNA polymerase mutants which can be used in the method of the invention
 CC of identifying and analysing polymorphic DNA. The method comprises
 CC reacting a DNA sample with at least one DNA polymerase having reduced
 CC ability to add non-templated nucleotides to the 3'-terminus and
 CC amplifying the fragment, followed by analysis, identification or typing
 CC of the amplicons. The DNA polymerases used are thermostable and the
 CC modification or mutation is in the O-helix of the DNA polymerase.
 CC Analysis of polymorphic DNA is particularly used to establish
 CC relationships between individuals (forensic or paternity testing),
 CC particularly where the polymorphism is a mini- or micro-satellite or
 CC short tandem repeat. More generally the DNA polymerase and its mutant
 CC form are useful in any process that involves nucleic acid amplification,
 CC e.g. identification of pathogens, cancer or genetic diseases (cystic
 CC fibrosis, haemophilia, Alzheimer's disease), screening organs or tissues
 CC before transplanting, diagnosis, plant breeding.
 CC Note: This sequence is not provided in the specification. It has been
 CC created by modifying the wild-type the DNA polymerase sequence

CC (AAW8758).
 XX
 SQ Sequence 609 AA;
 Query Match 95.9%; Score 47; DB 19; Length 609;
 Best Local Similarity 90.9%; Pred. No. 0.27;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVE 11
 ||||| ||||
 Db 457 LSVRLGIPYKE 467

RESULT 12
 AAW11514
 ID AAW11514 standard; Protein; 610 AA.
 XX
 AC AAW11514;
 XX
 DT 15-APR-1997 (first entry)
 XX
 DE Thermotoga neapolitana DNA polymerase The M284 mutant.
 XX
 KW DNA polymerase; The; thermostable enzyme; protein engineering;
 KM thermostability; DNA sequencing; amplification; PCR;
 KM polymerase chain reaction; exonuclease; Thermotoga neapolitana.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..14
 FT /label= 5'-Exonuclease_domain
 FT /note="Residues 1-14 comprise the C-terminal
 FT Domain 15..199
 FT /label= 3'-Exonuclease_domain
 FT Domain 200..610
 FT /label= Polymerisation_domain
 XX
 PN WO9641014-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US09641.
 XX
 PR 31-MAY-1996; 96US-0656664.
 PR 07-JUN-1995; 95US-0484661.
 XX
 PA (PROM-) PROMEXA CORP.
 XX
 PI Bolchakova E, Bartlett JK, Huang F, Miller KM, Otto P;
 PI Slater MR, Storis DR;
 XX
 DR WPI: 1997-052374/05.
 DR N-PADB; AAT58135.
 XX
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 XX
 PS Claim 11: Page 128-131; 200pp; English.
 XX
 CC DNA polymerase mutant The M284 (AAW11514) lacks the N-terminal 283
 CC amino acid residues of the native enzyme (AAW11513) of Thermotoga
 CC neapolitana. The naturally-occurring Met-284 residue becomes the
 CC initiating methionine. The mutant enzyme is obtd. by expression of
 CC DNA clone PM284 (AAT58135) produced by restriction enzyme digestion
 CC of the native gene (AAT58134). The N-terminal deletion eliminates
 CC the 5' exonuclease activity of the enzyme. The M284 has a half-
 CC life at 97.5 deg C of 18 min compared with 5 min for the native
 CC enzyme. Amino acid substns. in the 3' exonuclease domain of The
 CC M284 (see also AAW11516-22) yield enzymes esp. suitable for DNA
 CC sequencing and amplification reactions, e.g. PCR.

SQ Sequence 610 AA;
 Query Match 95.9%; Score 47; DB 18; Length 610;
 Best Local Similarity 90.9%; Pred. No. 0.27;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LSVRLGXPKK 11
 ||||| ||||
 DB 458 LSVRLGXPKK 468
 RESULT 13
 AAM11515
 ID AAM11515 standard; Protein: 610 AA.
 XX
 AC AAM11515;
 XX
 DT 15-APR-1997 (first entry)
 XX
 DE Thermotoga neapolitana DNA polymerase The M284(D323E) mutant.
 XX
 KM DNA polymerase; The; thermostable enzyme; protein engineering;
 KM thermostability; DNA sequencing; amplification; PCR;
 KM polymerase chain reaction; exonuclease; Thermotoga neapolitana.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 1..14
 FT Domain /label= 5'-Exonuclease_domain
 FT /note= "residues 1-14 comprise the C-terminal
 FT end of the The 5'-exonuclease domain"
 FT 15..199
 FT Domain /label= 3'-Exonuclease_domain
 FT /note= "the D323E substn. corresponds to amino
 FT acid position 40 in the mutant enzyme"
 FT 200..610
 FT Domain /label= Polymerisation_domain
 FT
 PN W09641014-A1.
 XX
 PD 19-DEC-1996.
 XX
 PE 07-JUN-1996; 96WO-US09641.
 XX
 PR 31-MAY-1996; 96US-0656664.
 PR 07-JUN-1995; 95US-0484661.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
 PI Slater MR, Storts DR;
 XX
 DR WPI: 1997-052374/05.
 DR N-PSDB; AAT58136.
 XX
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 XX
 PS Claim 11: Page 138-140; 200pp: English.
 XX
 CC DNA polymerase mutant The M284(D323E) (AAM11515) lacks the N-terminal
 CC 283 amino acid residues of the native enzyme (AAM11513) of Thermotoga
 CC neapolitana and contains a D323E substn. in the 3' exonuclease
 CC domain (numbering according to the native enzyme). It is obid. by
 CC expression of a DNA clone (AAT58136) created by site-directed
 CC mutagenesis of construct pM284 (see also AAT58135). The M284(D323E)
 CC has no significant 5' or 3' exonuclease activity and has a half-
 CC life at 97.5 deg C of 12.5 min compared with 5 min for native The
 CC and 8 min for native Tag. It can be used in DNA sequencing
 CC reactions and for the amplification of double stranded DNA.
 XX
 SO Sequence 610 AA;

Query Match 95.9%; Score 47; DB 18; Length 610;
 Best Local Similarity 90.9%; Pred. No. 0.27;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LSVRLGXPKK 11
 ||||| ||||
 DB 458 LSVRLGXPKK 468
 RESULT 14
 AAM11516
 ID AAM11516 standard; Protein: 610 AA.
 XX
 AC AAM11516;
 XX
 DT 15-APR-1997 (first entry)
 XX
 DE Thermotoga neapolitana DNA polymerase The M284(E325D) mutant.
 XX
 KM DNA polymerase; The; thermostable enzyme; protein engineering;
 KM thermostability; DNA sequencing; amplification; PCR;
 KM polymerase chain reaction; exonuclease; Thermotoga neapolitana.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 1..14
 FT Domain /label= 5'-Exonuclease_domain
 FT /note= "residues 1-14 comprise the C-terminal
 FT end of the The 5'-exonuclease domain"
 FT 15..199
 FT Domain /label= 3'-Exonuclease_domain
 FT /note= "the E325D substn. corresponds to amino
 FT acid position 42 in the mutant enzyme"
 FT 200..610
 FT Domain /label= Polymerisation_domain
 FT
 PN W09641014-A1.
 XX
 PD 19-DEC-1996.
 XX
 PE 07-JUN-1996; 96WO-US09641.
 XX
 PR 31-MAY-1996; 96US-0656664.
 PR 07-JUN-1995; 95US-0484661.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;
 PI Slater MR, Storts DR;
 XX
 DR WPI: 1997-052374/05.
 DR N-PSDB; AAT58137.
 XX
 PT Thermostable DNA polymerase from Thermotoga neapolitana - useful in
 PT DNA sequencing or for amplification of double stranded DNA
 XX
 PS Claim 11: Page 142-144; 200pp: English.
 XX
 CC DNA polymerase mutant The M284(E325D) (AAM11516) lacks the N-terminal
 CC 283 amino acid residues of the native enzyme (AAM11513) of Thermotoga
 CC neapolitana and contains a E325D substn. in the 3' exonuclease
 CC domain (numbering according to the native enzyme). It is obid. by
 CC expression of a DNA clone (AAT58137) created by site-directed
 CC mutagenesis of construct pM284 (see also AAT58135). The M284(E325D)
 CC has no significant 5' or 3' exonuclease activity and has a half-
 CC life at 97.5 deg C of 5 min (unchanged from the native enzyme).
 CC It can be used in DNA sequencing reactions and for the
 CC amplification of double stranded DNA.
 XX
 SO Sequence 610 AA;

Query Match 95.9%; Score 47; DB 18; Length 610;
 Best Local Similarity 90.9%; Pred. No. 0.27;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGIPVKE 11
 ||||| ||||
 Db 458 LSVRLGIPVKE 468

RESULT 15

AAW11517
 ID AAW11517 standard; Protein: 610 AA.

AC AAW11517;

DT 15-APR-1997 (first entry)

DE Thermotoga neapolitana DNA polymerase Tne M284(Y464F) mutant.

KW DNA polymerase; Tne; thermostable enzyme; protein engineering;

KW thermostability; DNA sequencing; amplification; PCR;

KW polymerase chain reaction; exonuclease; Thermotoga neapolitana.

OS Synthetic.

FH Key Location/Qualifiers

FT Domain 1..14

FT /label= 5'-Exonuclease_domain

FT /note= "residues 1-14 comprise the C-terminal

FT end of the Tne 5'-exonuclease domain"

FT Domain 15..199

FT /label= 3'-Exonuclease_domain

FT /note= "the Y464F substn. corresponds to amino

FT acid position 181 in the mutant enzyme"

FT Domain 200..610

FT /label= Polymerisation_domain

PN W09641014-A1.

PD 19-DEC-1996.

PF 07-JUN-1996; 96WO-US09641.

PR 31-MAY-1996; 96US-0656664.

PR 07-JUN-1995; 95US-0484661.

PA (PROM-) PROMEGA CORP.

PI Bolchakova E, Hartnett JR, Huang F, Miller KM, Otto P;

PI Slater MR, Storts DR;

DR WPI; 1997-052374/05.

DR N-PSDB; AAT58138.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 95.9%; Score 47; DB 18; Length 610;
 Best Local Similarity 90.9%; Pred. No. 0.27;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGIPVKE 11
 ||||| ||||
 Db 458 LSVRLGIPVKE 468

Search completed: January 15, 2003, 11:20:17
 Job time : 29.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:20:34 ; Search time 10.8571 Seconds

(Without alignments)
20.140 Million cell updates/sec

Title: US-09-823-649a-5

Perfect score: 49

Sequence: 1 LSVRLGXPKYKE 11

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCY_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCYUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	95.9	11 10 US-09-823-649A-5	Sequence 5, Appl1
2	47	95.9	11 10 US-09-823-649A-15	Sequence 15, Appl1
3	47	95.9	11 10 US-09-823-649A-16	Sequence 16, Appl1
4	47	95.9	610 9 US-09-891-332A-4	Sequence 4, Appl1
5	47	95.9	677 9 US-09-891-332A-3	Sequence 3, Appl1
6	47	95.9	708 9 US-09-891-332A-5	Sequence 5, Appl1
7	47	95.9	893 9 US-09-891-332A-2	Sequence 2, Appl1
8	47	95.9	893 9 US-09-891-332A-6	Sequence 6, Appl1
9	47	95.9	893 9 US-09-891-332A-7	Sequence 7, Appl1
10	47	95.9	893 9 US-09-891-332A-8	Sequence 8, Appl1
11	47	95.9	893 9 US-09-891-332A-9	Sequence 9, Appl1
12	47	95.9	893 10 US-09-741-664-1	Sequence 1, Appl1
13	47	95.9	893 10 US-09-741-664-2	Sequence 2, Appl1
14	47	95.9	11 12 US-10-053-632-4	Sequence 4, Appl1
15	47	95.9	11 12 US-10-053-632-5	Sequence 5, Appl1
16	47	95.9	11 12 US-10-053-632-6	Sequence 6, Appl1
17	47	95.9	11 12 US-10-053-632-15	Sequence 15, Appl1
18	47	95.9	11 12 US-10-053-632-16	Sequence 16, Appl1
19	47	95.9	11 12 US-10-053-632-17	Sequence 17, Appl1

ALIGNMENTS

20	41	83.7	11 12 US-10-052-417-4	Sequence 4, Appl1
21	41	83.7	11 12 US-10-052-417-5	Sequence 5, Appl1
22	41	83.7	11 12 US-10-052-417-6	Sequence 6, Appl1
23	41	83.7	11 12 US-10-052-417-15	Sequence 15, Appl1
24	41	83.7	11 12 US-10-052-417-16	Sequence 16, Appl1
25	41	83.7	11 12 US-10-052-417-17	Sequence 17, Appl1
26	41	83.7	610 9 US-09-891-332A-10	Sequence 10, Appl1
27	36	73.5	30 10 US-09-922-261-3	Sequence 34, Appl1
28	34	69.4	666 10 US-09-815-242-13637	Sequence 13637, A
29	34	69.4	676 10 US-09-815-242-13513	Sequence 13513, A
30	32	65.3	319 9 US-09-738-626-6900	Sequence 6900, Ap
31	31	63.3	96 10 US-09-867-550-952	Sequence 952, Ap
32	31	63.3	114 10 US-09-809-545A-74	Sequence 74, Appl
33	31	63.3	159 10 US-09-867-550-954	Sequence 954, Ap
34	31	63.3	189 10 US-09-216-993-130	Sequence 130, Appl
35	31	63.3	250 12 US-10-027-959-6	Sequence 6, Appl1
36	31	63.3	259 9 US-09-991-211-3	Sequence 3, Appl1
37	31	63.3	259 10 US-09-815-242-11603	Sequence 11603, A
38	30	61.2	95 9 US-09-895-913A-266	Sequence 266, App
39	30	61.2	102 9 US-10-001-857-189	Sequence 189, App
40	30	61.2	301 9 US-09-738-626-4278	Sequence 4278, Ap
41	30	61.2	380 10 US-09-815-242-11146	Sequence 11146, A
42	30	61.2	403 9 US-09-738-626-6217	Sequence 6217, Ap
43	30	61.2	408 10 US-09-761-640-6	Sequence 6, Appl1
44	30	61.2	471 9 US-09-955-732-21	Sequence 21, Appl1
45	30	61.2	471 10 US-09-761-640-4	Sequence 4, Appl1

RESULT 1

US-09-823-649A-5
Sequence 5, Application US-09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elstrom, Carla
APPLICANT: Gelfand, David
APPLICANT: Hinchell, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrunner, Nancy
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASE
FILE REFERENCE: RP10106
CURRENT APPLICATION NUMBER: US-09/823,649A
PRIOR FILING DATE: 2001-03-30
PRIORITY APPLICATION NUMBER: US 60/198,336
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 5
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: sequence motif
NAME/KEY: VARIANT
LOCATION: (7)..(7)
OTHER INFORMATION: X is V or I
US-09-823-649A-5

Query Match

Best Local Similarity 95.9%; Score 47; DB 10; Length 11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSVRLGXPKYKE 11

DB 1 LSVRLGXPKYKE 11

RESULT 2
US-09-823-649A-15

Sequence 15, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elfstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASES
FILE REFERENCE: RPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin version 3.0
SEQ ID NO 15
LENGTH: 11
TYPE: PRT
ORGANISM: Thermotoga maritima
US-09-823-649a-15

Query Match 95.9% Score 47; DB 10; Length 11;
Best Local Similarity 90.9% Pred. No. 0.00037;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVE 11
DB 1 LSVRLGXPKVE 11

RESULT 3
US-09-823-649a-16
Sequence 16, Application US/09823649A
Patent No. US20020012970A1
GENERAL INFORMATION:
APPLICANT: Smith, Edward
APPLICANT: Elfstrom, Carita
APPLICANT: Gelfand, David
APPLICANT: Higuchi, Russell
APPLICANT: Myers, Thomas
APPLICANT: Schoenbrunner, Nancy
APPLICANT: Wang, Alice
TITLE OF INVENTION: HIGH TEMPERATURE REVERSE TRANSCRIPTION USING MUTANT DNA POLYMERASES
FILE REFERENCE: RPA1006
CURRENT APPLICATION NUMBER: US/09/823,649A
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/198,336
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin version 3.0
SEQ ID NO 16
LENGTH: 11
TYPE: PRT
ORGANISM: Thermotoga neapolitana
US-09-823-649a-16

Query Match 95.9% Score 47; DB 10; Length 11;
Best Local Similarity 90.9% Pred. No. 0.00037;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVE 11
DB 1 LSVRLGXPKVE 11

RESULT 4
US-09-891-332A-4
Sequence 4, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:

APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shuwel
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
MOLECULAR TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-891-332A-4

Query Match 95.9% Score 47; DB 9; Length 610;
Best Local Similarity 90.9% Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKVE 11
DB 458 LSVRLGXPKVE 468

RESULT 5
US-09-891-332A-3
Sequence 3, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shuwel
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 677 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-891-332A-3

Query Match
Best Local Similarity 95.9%; Score 47; DB 9; Length 677;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRIGXPVK 11
||||| ||||
Db 525 LSVRIGXPVK 535

RESULT 6
US-09-891-332A-5
Sequence 5, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shweil
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 708 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-891-332A-5

Query Match
Best Local Similarity 95.9%; Score 47; DB 9; Length 708;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRIGXPVK 11
||||| ||||
Db 556 LSVRIGXPVK 566

RESULT 7
US-09-891-332A-2
Sequence 2, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shweil
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-891-332A-2

Query Match
95.9%; Score 47; DB 9; Length 893;

Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 LSVRLGXPKYE 11
Db 741 LSVRLGXPKYE 751

RESULT 8
US-09-891-332A-6
Sequence 6, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Sulus, Joseph
Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-891-332A-6
Query Match 95.9%; Score 47; DB 9; Length 893;
Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 LSVRLGXPKYE 11
Db 741 LSVRLGXPKYE 751

RESULT 9
US-09-891-332A-7
Sequence 7, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Sulus, Joseph
Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-891-332A-7
Query Match 95.9%; Score 47; DB 9; Length 893;
Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 LSVRLGXPKYE 11
Db 741 LSVRLGXPKYE 751

RESULT 10
US-09-891-332A-8
Sequence 8, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Sulus, Joseph
Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SOURCE DESCRIPTION: SEQ ID NO: 8:
US-09-891-332A-8
Query Match
Best Local Similarity 90.9%; Score 47; DB 9; Length 893;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGIPVKE 11
DB 741 LSVRLGIPVKE 751
RESULT 11
US-09-891-332A-9
Sequence 9, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SOURCE DESCRIPTION: SEQ ID NO: 9:
US-09-891-332A-9
Query Match
Best Local Similarity 90.9%; Score 47; DB 9; Length 893;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGIPVKE 11
DB 741 LSVRLGIPVKE 751
RESULT 12
US-09-741-664-1
Sequence 1, Application US/09741664
Patent No. US2001004134A1
GENERAL INFORMATION:
APPLICANT: Rashid, Ayoub
Solus, Joseph
TITLE OF INVENTION: Stable Compositions for Nucleic Acid
Sequencing and Amplification
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/741,664
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/049,021
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3910000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-741-664-1
Query Match
Best Local Similarity 90.9%; Score 47; DB 10; Length 893;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGIPVKE 11
DB 741 LSVRLGIPVKE 751

```
RESULT 13
US-09-741-664-2
; Sequence 2, Application US/09741664
; Patent No. US20010041334A1
; GENERAL INFORMATION:
; APPLICANT: Rasthchian, Ayoub
; TITLE OF INVENTION: Stable Compositions for Nucleic Acid
; TITLE OF INVENTION: Sequencing and Amplification
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/741,664
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/049,021
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmoud, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3910000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 893 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-741-664-2
Query Match 95.9%; Score 47; DB 10; Length 893;
Best Local Similarity 90.9%; Pred. No. 0.052;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 LSVRLGXPKVE 11
|||||
DB 741 LSVRLGXPKVE 751

RESULT 14
US-10-053-632-4
; Sequence 4, Application US/10053632
; Patent No. US2002095560A1
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David
; APPLICANT: Kalman, Lisa
; APPLICANT: Reichert, Fred
; APPLICANT: Signa, Chris
; APPLICANT: Myers, Thomas
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES INCORPORATING NUCLEOTIDE TRIPHOSPHAT
; TITLE OF INVENTION: LABELED WITH FLUORESCENT DYES
; FILE REFERENCE: 1803-329-999
; CURRENT APPLICATION NUMBER: US/10/053,632
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
```

```
; SEQ ID NO 4
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence from mutant
; OTHER INFORMATION: thermostable DNA polymerase
; NAME/KEY: VARIANT
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa at position 4 is any amino acid except glutamic acid resid
; NAME/KEY: VARIANT
; LOCATION: (7)..(7)
; OTHER INFORMATION: Xaa at position 7 is Val or Ile
US-10-053-632-4
Query Match 83.7%; Score 41; DB 12; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0068;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Caps 0;
```

```
RESULT 15
US-10-053-632-5
; Sequence 5, Application US/10053632
; Patent No. US2002095560A1
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David
; APPLICANT: Kalman, Lisa
; APPLICANT: Reichert, Fred
; APPLICANT: Signa, Chris
; APPLICANT: Myers, Thomas
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES INCORPORATING NUCLEOTIDE TRIPHOSP
; TITLE OF INVENTION: LABELED WITH FLUORESCENT FAMILY DYES
; FILE REFERENCE: 1803-329-999
; CURRENT APPLICATION NUMBER: US/10/053,632
; CURRENT FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence from mutant
; OTHER INFORMATION: thermostable DNA polymerase
; NAME/KEY: VARIANT
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa at position 4 is any amino acid
US-10-053-632-5
Query Match 83.7%; Score 41; DB 12; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0068;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Caps 0;
```

```
QY 1 LSVRLGXPKVE 11
|||||
DB 1 LSVRLGXPKVE 11
```

Search completed: January 15, 2003, 12:39:51
Job time : 10.8571 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2003, 11:17:28 ; Search time 9.28571 Seconds
(without alignments)
34.855 Million cell updates/sec

Title: US-09-823-649A-5
Perfect score: 49
Sequence: 1 LSVRLGXPVKE 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5a_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5a_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/5a_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/5a_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/5a_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/5a_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	95.9	230	2	US-08-537-400-16
2	47	95.9	230	2	US-08-706-702-18
3	47	95.9	230	3	US-08-706-706-18
4	47	95.9	494	3	US-08-484-661A-39
5	47	95.9	494	3	US-08-656-664-39
6	47	95.9	494	5	PCT-US96-09641-39
7	47	95.9	571	3	US-08-484-661A-37
8	47	95.9	571	3	US-08-656-664-37
9	47	95.9	571	5	PCT-US96-09641-37
10	47	95.9	578	3	US-08-484-661A-11
11	47	95.9	578	3	US-08-656-664-11
12	47	95.9	578	5	PCT-US96-09641-11
13	47	95.9	610	3	US-08-484-661A-8
14	47	95.9	610	3	US-08-484-661A-16
15	47	95.9	610	3	US-08-484-661A-19
16	47	95.9	610	3	US-08-484-661A-23
17	47	95.9	610	3	US-08-484-661A-26
18	47	95.9	610	3	US-08-484-661A-29
19	47	95.9	610	3	US-08-484-661A-33
20	47	95.9	610	3	US-08-484-661A-35
21	47	95.9	610	3	US-08-656-664-8
22	47	95.9	610	3	US-08-656-664-16
23	47	95.9	610	3	US-08-656-664-19
24	47	95.9	610	3	US-08-656-664-23
25	47	95.9	610	3	US-08-656-664-26
26	47	95.9	610	3	US-08-656-664-29
27	47	95.9	610	3	US-08-656-664-33

28	47	95.9	610	3	US-08-656-664-35	Sequence 35, Appl
29	47	95.9	610	3	US-08-656-664-54	Sequence 54, Appl
30	47	95.9	610	4	US-09-019-160-4	Sequence 4, Appl
31	47	95.9	610	5	PCT-US96-09641-8	Sequence 8, Appl
32	47	95.9	610	5	PCT-US96-09641-16	Sequence 16, Appl
33	47	95.9	610	5	PCT-US96-09641-19	Sequence 19, Appl
34	47	95.9	610	5	PCT-US96-09641-23	Sequence 23, Appl
35	47	95.9	610	5	PCT-US96-09641-26	Sequence 26, Appl
36	47	95.9	610	5	PCT-US96-09641-29	Sequence 29, Appl
37	47	95.9	610	5	PCT-US96-09641-33	Sequence 33, Appl
38	47	95.9	610	5	PCT-US96-09641-35	Sequence 35, Appl
39	47	95.9	610	5	PCT-US96-09641-34	Sequence 34, Appl
40	47	95.9	677	4	US-09-019-160-3	Sequence 3, Appl
41	47	95.9	708	4	US-09-019-160-5	Sequence 5, Appl
42	47	95.9	893	1	US-07-977-434-4	Sequence 4, Appl
43	47	95.9	893	1	US-08-458-819-4	Sequence 4, Appl
44	47	95.9	893	2	US-08-706-702-3	Sequence 3, Appl
45	47	95.9	893	3	US-08-484-661A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-537-400-16
; Sequence 16, Application US/08537400
; Patent No. 5939301
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Cloned DNA Polymerases From Thermotoga
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,400
; FILING DATE: 02-OCT-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,423
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/370,190
; FILING DATE: 09-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-537-400-16
Query Match 95.9% Score 47; DB 2; Length 230;
Best Local Similarity 90.9% Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0;

OY 1 LSVRLGXPKYKE 11
||||| ||||
DB 78 LSVRLGXPKYKE 88

RESULT 2
US-08-706-702-18
; Sequence 18, Application US/08706702
; Patent No. 5948614

GENERAL INFORMATION:
; APPLICANT: Chatterjee, Deb K.
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga
; TITLE OF INVENTION: maritima and Mutants Thereof
; NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,702
; FILING DATE: 06-SEP-1996

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/689,807
; FILING DATE: 14-AUG-1996

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,400
; FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/576,759
; FILING DATE: 21-DEC-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,397
; FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/525,057
; FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800006

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-706-702-18
Query Match 95.9%; Score 47; DB 2; Length 230;
Best local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGXPKYKE 11
||||| ||||
DB 78 LSVRLGXPKYKE 88

RESULT 3
US-08-706-706-18
; Sequence 18, Application US/08706706
; Patent No. 6015668

GENERAL INFORMATION:

APPLICANT: Hughes, A. John
; APPLICANT: Chatterjee, Deb K.
; TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga and
; NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 New York Ave., N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,706
; FILING DATE: 06-SEP-1996

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/689,807
; FILING DATE: 14-AUG-1996

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,400
; FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/370,190
; FILING DATE: 01-JAN-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,423
; FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/576,759
; FILING DATE: 21-DEC-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/537,397
; FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/525,057
; FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.2800005

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-706-706-18
Query Match 95.9%; Score 47; DB 3; Length 230;
Best local Similarity 90.9%; Pred. No. 0.012;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGXPKYKE 11
||||| ||||
DB 78 LSVRLGXPKYKE 88

RESULT 4
US-08-484-661A-39

; Sequence 39, Application US/08484661A
; Patent No. 6001645
; GENERAL INFORMATION:
; APPLICANT: SLATER, MICHAEL R.
; APPLICANT: HARTNETT, JAMES R.

```

?
? APPLICANT: HUANG, FEN
? APPLICANT: BOLCHAKOVA, ELENA
? TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
? TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
? NUMBER OF SEQUENCES: 51
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: MEDLEN & CARROLL
? STREET: 220 MONTGOMERY STREET, SUITE 2200
? CITY: SAN FRANCISCO
? STATE: CALIFORNIA
? COUNTRY: UNITED STATES OF AMERICA
? ZIP: 94104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/484,661A
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: INGOLIA, DIANE E.
? REGISTRATION NUMBER: 40,027
? REFERENCE/DOCKET NUMBER: PRNG-01175
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 39:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 494 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-484-661A-39

Query Match          95.9%; Score 47; DB 3; Length 494;
Best Local Similarity 90.9%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKK 11
Db 342 LSVRLGXPKK 352

RESULT 5
US-08-656-664-39
? Sequence 39, Application US/0865664
? Patent No. 6077664
? GENERAL INFORMATION:
? APPLICANT: Slater, Michael R.
? APPLICANT: Huang, Fen
? APPLICANT: Hartnett, James R.
? APPLICANT: Bolchakova, Elena
? APPLICANT: Storts, Douglas R.
? APPLICANT: Otto, Paul
? TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
? TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
? NUMBER OF SEQUENCES: 57
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Medlen & Carroll
? STREET: 220 Montgomery Street, Suite 2200
? CITY: San Francisco
? STATE: California
? COUNTRY: United States Of America
? ZIP: 94104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/656,664
```

```

?
? FILING DATE: 31-MAY-1996
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: INGOLIA, DIANE E.
? REGISTRATION NUMBER: 40,027
? REFERENCE/DOCKET NUMBER: PRNG-02185
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 39:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 494 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-656-664-39

Query Match          95.9%; Score 47; DB 3; Length 494;
Best Local Similarity 90.9%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPKK 11
Db 342 LSVRLGXPKK 352

RESULT 6
PCT-US96-09641-39
? Sequence 39, Application PCT/US9609641
? GENERAL INFORMATION:
? APPLICANT: Slater, Michael R.
? APPLICANT: Huang, Fen
? APPLICANT: Hartnett, James R.
? APPLICANT: Bolchakova, Elena
? APPLICANT: Storts, Douglas R.
? APPLICANT: Otto, Paul
? TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
? TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
? NUMBER OF SEQUENCES: 57
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Medlen & Carroll
? STREET: 220 Montgomery Street, Suite 2200
? CITY: San Francisco
? STATE: California
? COUNTRY: United States Of America
? ZIP: 94104
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US96/09641
? FILING DATE: 31-MAY-1996
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: INGOLIA, DIANE E.
? REGISTRATION NUMBER: 40,027
? REFERENCE/DOCKET NUMBER: PRNG-02185
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 705-8410
? TELEFAX: (415) 397-8338
? INFORMATION FOR SEQ ID NO: 39:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 494 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? PCT-US96-09641-39

Query Match          95.9%; Score 47; DB 5; Length 494;
Best Local Similarity 90.9%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 LSVRLGXPKVE 11
|||||
Db 342 LSVRLGXPKVE 352

RESULT 7
US-08-484-661A-37
; Sequence 37, Application US/08484661A
; Patent No. 6001645
; GENERAL INFORMATION:
; APPLICANT: SLATER, MICHAEL R.
; APPLICANT: HARTNETT, JAMES R.
; APPLICANT: HUANG, FEN
; APPLICANT: BOLCHAKOVA, ELENA
; TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
; TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,661A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: INGOLIA, DIANE E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: PMG-01175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-661A-37
Query Match 95.9%; Score 47; DB 3; Length 571;
Best Local Similarity 90.9%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGXPKVE 11
|||||
Db 419 LSVRLGXPKVE 429

RESULT 8
US-08-656-664-37
; Sequence 37, Application US/08656664
; Patent No. 6077664
; GENERAL INFORMATION:
; APPLICANT: Slater, Michael R.
; APPLICANT: Huang, Fen
; APPLICANT: Hartnett, James R.
; APPLICANT: Bolchakova, Elena
; APPLICANT: Storts, Douglas R.
; APPLICANT: Otto, Paul
; TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
; TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,664
; FILING DATE: 31-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: PMG-02185
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-656-664-37
Query Match 95.9%; Score 47; DB 3; Length 571;
Best Local Similarity 90.9%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LSVRLGXPKVE 11
|||||
Db 419 LSVRLGXPKVE 429

RESULT 9
PCT-US96-09641-37
; Sequence 37, Application PCT/US9609641
; GENERAL INFORMATION:
; APPLICANT: Slater, Michael R.
; APPLICANT: Huang, Fen
; APPLICANT: Hartnett, James R.
; APPLICANT: Bolchakova, Elena
; APPLICANT: Storts, Douglas R.
; APPLICANT: Otto, Paul
; TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
; TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09641
; FILING DATE: 31-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: PMG-02185
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 571 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-09641-37

Query Match 95.9%; Score 47; DB 5; Length 571;
Best Local Similarity 90.9%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 11
|||||
DB 419 LSVRLGIPVKE 429

RESULT 10
US-08-484-661A-11
Sequence 11, Application US/08484661A
Patent No. 6001645
GENERAL INFORMATION:
APPLICANT: SLATER, MICHAEL R.
APPLICANT: HARTNETT, JAMES R.
APPLICANT: HUANG, FEN
APPLICANT: BOLCHAKOVA, ELENA
TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,661A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-01175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 578 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-661A-11

Query Match 95.9%; Score 47; DB 3; Length 578;
Best Local Similarity 90.9%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 11
|||||
DB 426 LSVRLGIPVKE 436

RESULT 11
US-08-656-664-11

Sequence 11, Application US/08656664
Patent No. 6077664
GENERAL INFORMATION:
APPLICANT: Slater, Michael R.
APPLICANT: Huang, Fen
APPLICANT: Hartnett, James R.
APPLICANT: Bolchakova, Elena
APPLICANT: Storts, Douglas R.
APPLICANT: Otto, Paul
TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,664
FILING DATE: 31-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-02185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 578 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-656-664-11

Query Match 95.9%; Score 47; DB 3; Length 578;
Best Local Similarity 90.9%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGXPVKE 11
|||||
DB 426 LSVRLGIPVKE 436

RESULT 12
PCT-US96-09641-11
Sequence 11, Application PC/TUS9619641
GENERAL INFORMATION:
APPLICANT: Slater, Michael R.
APPLICANT: Huang, Fen
APPLICANT: Hartnett, James R.
APPLICANT: Bolchakova, Elena
APPLICANT: Storts, Douglas R.
APPLICANT: Otto, Paul
TITLE OF INVENTION: THERMOPHILIC DNA POLYMERASES FROM
TITLE OF INVENTION: THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/09641
FILING DATE: 31-MAY-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-02185
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 578 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-09641-11

Query Match
Best Local Similarity 95.9%; Score 47; DB 5; Length 578;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGXPYKE 11
||||| |||
DB 426 LSVRLGIPYKE 436

RESULT 13
US-08-484-661A-8
Sequence 8, Application US/08484661A
Patent No. 6001645
GENERAL INFORMATION:
APPLICANT: SLATER, MICHAEL R.
APPLICANT: HARTNETT, JAMES R.
APPLICANT: HUANG, FEN
APPLICANT: BOLCHAKOVA, ELENA
TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,661A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-01175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-484-661A-8
Query Match
Best Local Similarity 95.9%; Score 47; DB 3; Length 610;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGXPYKE 11
||||| |||
DB 458 LSVRLGIPYKE 468

RESULT 14
US-08-484-661A-16
Sequence 16, Application US/08484661A
Patent No. 6001645
GENERAL INFORMATION:
APPLICANT: SLATER, MICHAEL R.
APPLICANT: HARTNETT, JAMES R.
APPLICANT: HUANG, FEN
APPLICANT: BOLCHAKOVA, ELENA
TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES
TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,661A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: PRMG-01175
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-661A-16

Query Match
Best Local Similarity 95.9%; Score 47; DB 3; Length 610;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LSVRLGXPYKE 11
||||| |||
DB 458 LSVRLGIPYKE 468

RESULT 15
US-08-484-661A-19
Sequence 19, Application US/08484661A
Patent No. 6001645
GENERAL INFORMATION:
APPLICANT: SLATER, MICHAEL R.
APPLICANT: HARTNETT, JAMES R.
APPLICANT: HUANG, FEN
APPLICANT: BOLCHAKOVA, ELENA
TITLE OF INVENTION: MODIFIED THERMOPHILIC DNA POLYMERASES

;; TITLE OF INVENTION: FROM THERMOTOGA NEAPOLITANA
;; NUMBER OF SEQUENCES: 51
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MEDLEN & CARROLL
;; STREET: 220 MONTGOMERY STREET, SUITE 2200
;; CITY: SAN FRANCISCO
;; STATE: CALIFORNIA
;; COUNTRY: UNITED STATES OF AMERICA
;; ZIP: 94104
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/484,661A
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: INGOLIA, DIANE E.
;; REGISTRATION NUMBER: 40,027
;; REFERENCE/DOCKET NUMBER: PRWG-01175
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;; INFORMATION FOR SEQ ID NO: 19:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 610 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-484-661A-19

Query Match 95.9%; Score 47; DB 3; length 610;
Best Local Similarity 90.9%; Pred. No. 0.037;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSVRLGAPVKE 11
|||||
Db 458 LSVRLGIPVKE 468

Search completed: January 15, 2003, 12:38:24
Job time : 10.2857 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame-plus.p2n model

Run on: January 15, 2003, 11:21:44 ; Search time 130.143 Seconds
(without alignments)
190.345 Million cell updates/sec

Title: US-09-823-649a-4
Perfect score: 50
Sequence: 1 LSXELSTPYEE 11

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: --
-MODE=frame+ .p2n.model -DEV=xlh
-O=/cgn2.1/USPRO.spool/US09823649.tunat.14012003.151001.29119/app.query.fasta.1.1393
-DB=N_Geneseq.101002 -OPMT=firstap -SUFFIX=ring -MINMATCH=0.1 -LOOPTC=0
-LOOPTX=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=plto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09823649.@cgn2.1.0.@runat.14012003.151001.29119 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEOUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -XGAPOP=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB	ID	Description
1	48	96.0	1635	AAQ24010	Mutant thermostabl
2	48	96.0	1899	AAQ24009	Mutant thermostabl
3	48	96.0	2043	AAQ24008	Mutant thermostabl
4	48	96.0	2277	AAQ24007	Mutant thermostabl
5	48	96.0	2367	AAQ24006	Mutant thermostabl
6	48	96.0	2493	AAQ24005	Mutant thermostabl
7	48	96.0	2495	AAQ80750	T. flavus DNA-poly
8	48	96.0	2496	AAQ80750	Thermus flavus DNA
9	48	96.0	2496	AAV65780	DNA sequence encod
10	48	96.0	2512	AAQ24369	Tfil DNA-polymeras
11	48	96.0	4947	AAQ86927	Tsp317 polymerase
12	45	90.0	1600	AAQ80745	5' Nuclease from T
13	45	90.0	1600	AAV27646	Mutant Thermus aqu
14	45	90.0	1600	AAV70347	Synthesis deficient
15	45	90.0	1600	AAV76647	Tag gene 5' nuclea
16	45	90.0	1600	AAV65786	Thermus aquaticus
17	45	90.0	1600	AAV53855	Nucleotide sequenc
18	45	90.0	1600	AAV63407	DNA sequence of a
19	45	90.0	1635	AAQ23998	Mutant thermostabl
20	45	90.0	1635	AAQ24322	Mutant thermostabl
21	45	90.0	1635	AAQ24328	Mutant thermostabl
22	45	90.0	1682	AAZ29460	Thermostable DNA d
23	45	90.0	1686	AAV47959	Mutant Tag polymer
24	45	90.0	1686	AAV47961	Mutant Tag polymer
25	45	90.0	1686	AAV27130	Py2 polymerase cod
26	45	90.0	1686	AAV27132	Thermostable DNA d
27	45	90.0	1689	AAV47960	Thermus flavus DNA
28	45	90.0	1689	AAV47961	DNA polymerase I e
29	45	90.0	1689	AAV27131	Chimeric Caurtag DN
30	45	90.0	1794	AAV32327	Mutant thermostabl
31	45	90.0	1794	AAV27255	Mutant thermostabl
32	45	90.0	1812	AAV00717	Thermus flavus DNA
33	45	90.0	1812	AAV00718	DNA polymerase I e
34	45	90.0	1886	AAQ23997	Chimeric Caurtag DN
35	45	90.0	1889	AAQ24321	Mutant thermostabl
36	45	90.0	1899	AAQ24327	Mutant thermostabl
37	45	90.0	1899	AAQ24819	Mutant thermostabl
38	45	90.0	1904	AAQ24832	Sulfolobus solifara
39	45	90.0	1965	AAQ24833	Sulfolobus acidoca
40	45	90.0	2043	AAQ23996	Polylysine (Pl) -
41	45	90.0	2043	AAQ23997	Mutant thermostabl
42	45	90.0	2043	AAQ24320	Mutant thermostabl
43	45	90.0	2274	AAQ24326	Mutant thermostabl
44	45	90.0	2277	AAQ24994	Mutant thermostabl
45	45	90.0	2277	AAQ24013	Mutant thermostabl
				AAQ24325	Mutant thermostabl

ALIGNMENTS

RESULT 1
AAQ24010
AAQ24010 standard: DNA: 1635 BP.

22-OCT-1992 (first entry)

Mutant thermostable DNA polymerase pspSA288.

5'-3': exonuclease; PCR; amplification; SSR; sequencing; PCR; ss.

Thermus species sps17.

Key old_sequence Location/Qualifiers

FT 1..2 /tag= a
FT /note= "nucleotides 4-861 deleted from the native
FT